



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup> :</b> C12N 9/16, 15/00, C12Q 1/68 C12Q 1/42, A61K 39/00, 37/54	<b>A1</b>	<b>(11) International Publication Number:</b> WO 93/18139 <b>(43) International Publication Date:</b> 16 September 1993 (16.09.93)
<b>(21) International Application Number:</b> PCT/US93/02172 <b>(22) International Filing Date:</b> 10 March 1993 (10.03.93)  <b>(30) Priority data:</b> 07/849,219                      10 March 1992 (10.03.92)                      US  <b>(71) Applicant:</b> LA JOLLA CANCER RESEARCH FOUNDATION [US/US]; 10901 North Torrey Pines Road, La Jolla, CA 92037 (US).  <b>(72) Inventor:</b> MILLAN, Jose, L. ; 10858 Caminito Alto, San Diego, CA 92131 (US).  <b>(74) Agents:</b> CAMPBELL, Cathryn et al.; Campbell & Flores, 4370 La Jolla Village Drive, Suite 700, San Diego, CA 92122 (US).		<b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i> <i>With amended claims and statement.</i>
<b>(54) Title:</b> RECOMBINANT CALF INTESTINAL ALKALINE PHOSPHATASE  <b>(57) Abstract</b>  The invention relates to isolated nucleic acids encoding recombinant calf intestinal alkaline phosphatase. Expression vectors and host cells transformed or transfected with such vectors are also provided. The invention further provides multifunctional polypeptides containing amino acid sequences encoding for calf intestinal alkaline phosphatase and a second amino acid sequence encoding a reagent having specific reactivity with a ligand. The recombinant calf intestinal alkaline phosphatase or its active fragments and the multifunctional polypeptides can be used in the methods for determining the presence or concentration of a ligand.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SK	Slovak Republic
CI	Côte d'Ivoire	LJ	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	MC	Monaco	TG	Togo
DE	Germany	MG	Madagascar	UA	Ukraine
DK	Denmark	ML	Mali	US	United States of America
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

## RECOMBINANT CALF INTESTINAL ALKALINE PHOSPHATASE

BACKGROUND OF THE INVENTION

The present invention relates to recombinant calf intestinal alkaline phosphatase and more particularly to isolated nucleic acids encoding the recombinant form of calf intestinal alkaline phosphatase.

Alkaline phosphatases (APs) are a family of functionally related enzymes named after the tissues in which they predominately appear. Such enzymes carry out hydrolase/transferase reactions on phosphate-containing substrates at a high pH optimum. The exact role of APs in biological processes remains poorly defined.

In humans and other higher animals, the AP family contains four members that are each encoded by a separate gene locus as reviewed in Millan, Anticancer Res. 8:995-1004 (1988) and Harris, Clin. Chem. Acta 186:133-150 (1989). The alkaline phosphatase family includes the tissue specific APs (placental AP, germ cell AP and intestinal AP) and the tissue non-specific AP found predominately in the liver, bone and kidney.

Intestinal alkaline phosphatase (IAP) derived from humans has been extensively characterized. As with all known APs, human IAP appears as a dimer, which is referred to as p75/150 in Latham & Stanbridge, P.N.A.S. (USA) 87:1263-1267 (1990). A cDNA clone for human adult IAP has been isolated from a  $\lambda$ gt11 expression library. This cDNA clone is 2513 base pairs in length and contains an open reading frame that encodes a 528 amino acid polypeptide as described in Henthorn et al., P.N.A.S. (USA) 84:1234-1238 (1987). IAP has also been found in other species, such as mice, cows, and fish as reported in McComb et al., Alkaline Phosphatases (Plenum, New York, 1989).

Generally, alkaline phosphatases are useful diagnostically in liver and bone disorders as described in McComb et al., supra, or for certain cancers as reviewed in Millan, Prog. Clin. Biol. Res., 344:453-475 (1990). APs are also useful as reagents in molecular biology. Of the known APs, bovine IAP has the highest catalytic activity. This property has made bovine IAP highly desirable for such biotechnological applications as enzyme-conjugates for use as diagnostics reagents or dephosphorylation of DNA, for example.

The isozymes of bovine IAP (b.IAP), including calf IAP, adult bovine IAP, and a tissue non-specific isozyme extracted from the small intestines, have been characterized by Besman & Coleman, J. Biol. Chem., 260:1190-1193 (1985). Although it is possible to purify naturally-occurring calf IAP extracted from intestinal tissues, it is technically very difficult to obtain an enzyme preparation of reproducible quality and purity. Generally, the enzymes are extracted from bovine intestines obtained from slaughter houses. Since the sacrificed animals are not of the same age, the proportion of the known b.IAP isozymes will vary significantly among the purified extracts.

Moreover, the intestine is known to contain high amounts of peptidases and glycosidases that degrade the naturally occurring IAP. Since the time from slaughter to enzyme extraction varies greatly, the amount of degradation will also vary greatly, resulting in a mixture of intact and several degradation products. Accordingly, the known methods of purifying IAP from naturally-occurring sources produce microheterogeneity in the purified IAP preparations. These partially degraded IAP molecules are technically difficult to separate from the native intact IAP molecules.

Due in part to the technical problems of separating intact b.IAP from degraded or partially processed calf IAP and the minute quantities of purified intact b.IAP that can be obtained from naturally-occurring sources, it has been difficult to determine the amino acid sequence encoding calf IAP. In addition, attempts to crystallize the IAP protein to determine the three-dimensional structure from the natural source has been hampered because of such microheterogeneity of the enzyme obtained from natural sources. It has only been possible to obtain small crystals of the natural enzyme, which are of insufficient quality for crystallographic studies.

Thus, a need exists for a homogeneous source of calf intestine alkaline phosphatase. Such a source would ideally provide an ample supply of pure, intact calf IAP for research and commercial use without time-consuming and labor intensive procedures. The present invention satisfies this need and provides related advantages as well.

20

#### SUMMARY OF THE INVENTION

The present invention generally relates to recombinant calf intestinal alkaline phosphatase (calf IAP) having an amino acid sequence substantially the same as naturally occurring calf IAP or its active fragments. The invention further provides isolated nucleic acids encoding such polypeptides. Vectors containing these nucleic acids and recombinant host cells transformed or transfected with such vectors are also provided.

Nucleic acid probes having nucleotide sequences complementary to a portion of the nucleotide sequence encoding calf IAP are also provided. Such probes can be used for the detection of nucleic acids encoding calf IAP or active fragments thereof.

The present invention further provides a multifunctional polypeptide containing an amino acid sequence of calf IAP and a second amino acid sequence having specific reactivity with a desired ligand. The  
5 second amino acid sequence can encode, for example, an antibody sequence when the desired ligand is an antigen.

The pure recombinant polypeptides of the present invention, including the multifunctional polypeptides, are particularly useful in methods for detecting the presence  
10 of antigens or other ligands in substances, such as fluid samples and tissues. Such diagnostic methods can be used for in vitro detection of such ligands.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 (SEQ ID NO: 9) shows the full length  
15 genomic sequence of calf IAP and the deduced amino acid sequence.

Figure 2 shows the restriction map of the entire calf IAP gene and the full length cDNA.

Figure 3 (SEQ ID NOS: 10-13) shows a comparison  
20 of IAPs from calf (b.IAP), rat (r.IAP), mouse (m.IAP), and human (h.IAP).

Figure 4 shows the results of studies relating to the heat inactivation of purified and recombinant calf IAP.

#### DETAILED DESCRIPTION OF THE INVENTION

25 The present invention relates to the elucidation of the calf intestinal alkaline phosphatase gene. More specifically, the invention relates to the nucleotide sequence of the region of the gene encoding the enzyme.



Previous attempts to produce a full length cDNA or a complete genomic clone for calf IAP have been unsuccessful. RNA extracted from bovine intestinal tissues are not fully processed (i.e., incompletely spliced RNA) or are quickly degraded after death. As such, only fragments of the genome coding region could be obtained.

It was through the extensive experimentation as set forth in the examples below that the full length cDNA clone of calf IAP was determined. Accordingly, the present invention is directed to isolated nucleic acids comprising the nucleotide sequence encoding calf IAP or an active fragment thereof having the enzymatic activity of the intact calf IAP. The nucleic acids can be DNA, cDNA or RNA.

The nucleic acid can have the nucleotide sequence substantially the same as the sequence identified in Figure 1, which shows the complete coding region of the genomic sequence of calf IAP. This nucleic acid (5.4 kb) contains 11 exons separated by 10 small introns at positions identical to those of other members of the tissue-specific AP family. Additionally, a 1.5 kb of the 5' sequence contains putative regulatory elements having homology to human and mouse IAP promoter sequences.

As used herein, the term "substantially the sequence" means the described nucleotide or amino acid sequence or other sequences having one or more additions, deletions or substitutions that do not substantially affect the ability of the sequence to encode a polypeptide having a desired activity, such as calf IAP or its active fragments. Thus, modifications that do not destroy the encoded enzymatic activity are contemplated.

As used herein, an active fragment of calf IAP refers to portions of the intact enzyme that substantially

retains the enzymatic activity of the intact enzyme. The retention of activity can be readily determined using methods known to those skilled in the art.

The terms "isolated" and "substantially purified" are used interchangeably and mean the polypeptide or nucleic acid is essentially free of other biochemical moieties with which it is normally associated in nature. Recombinant polypeptides are generally considered to be substantially purified.

The present invention further relates to expression vectors into which the coding region of the calf IAP gene can be subcloned. "Vectors" as used herein are capable of expressing nucleic acid sequences when such sequences are operationally lined to other sequences capable of effecting their expression. These expression vectors must be replicable in the host organisms either as episomes or as an integral part of the chromosomal DNA. Lack of replicability would render them effectively inoperable. In general, useful vectors in recombinant DNA techniques are often in the form of plasmids, which refer to circular, double stranded DNA loops which are not bound to the chromosome in their vector form. Suitable expression vectors can be plasmids such as, for example, pcDNA1 (Invitrogen, San Diego, CA).

A number of procaryotic expression vectors are known in the art, such as those disclosed, for example, in U.S. Patent Nos. 4,440,859; 4,436,815; 4,431,740; 4,431,739; 4,428,941; 4,425,437; 4,418,149; 4,411,994 and 4,342,832, all incorporated herein by reference. Eucaryotic systems and yeast expression vectors can also be used as described, for example, in U.S. Patent Nos. 4,446,235; 4,443,539; and 4,430,428, all incorporated herein by reference.



The vectors can be used to transfect or transform suitable host cells by various methods known in the art, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY (1989). Such  
5 host cells can be either eucaryotic or procaryotic cells. Examples of such hosts include chinese hamster ovary (CHO) cells, E.Coli and baculovirus infected insect cells. As used herein, "host cells" or "recombinant host cells" refer not only to the particular subject cell but to the progeny  
10 or potential progeny of such cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

15 The present invention further relates to recombinant proteins or polypeptides produced by the recombinant host cells of the present invention. The recombinant calf IAP protein has been characterized in terms of its heat stability up to about 50°C,  
20 electrophoretic and isoelectric focusing (IEF) behavior and kinetic parameters. The recombinant calf IAP protein of the present invention demonstrated displayed kinetic properties comparable to commercially available purified calf IAP, while showing less heterogeneity than the  
25 commercial enzymes in polyacrylamide gel electrophoresis and IEF, as described in the examples below.

Methods for obtaining or isolating recombinant calf IAP or active fragments are also provided. Such methods include culturing the recombinant host cells in a  
30 suitable growth medium. The protein or active fragments can thereafter be isolated from the cells by methods known in the art. If the expression system secretes calf IAP protein into growth media, the protein can be purified directly from cell-free media. If the protein is not  
35 secreted, it can be isolated from cell lysates. The

selection of the appropriate growth conditions and recovery methods are within the knowledge of one skilled in the art. Recombinant calf IAP or active fragments thereof can be unglycosylated or have a different glycosylation pattern  
5 than the native enzyme depending on the host that is used to prepare it.

The present invention further provides isolated nucleic acids containing a nucleotide sequence encoding calf IAP or an active fragment thereof and a second  
10 nucleotide sequence encoding a polypeptide having specific reactivity with a ligand. Such nucleic acids encode a chimeric or multifunctional polypeptide in which a region of the polypeptide has enzymatic activity conferred by the calf IAP sequence attached to a second region having  
15 specific reactivity with a particular ligand. Such multifunctional polypeptides are particularly useful in diagnostic assays for determining the presence or concentration of a particular ligand in a sample. The ligand can be, for example, a cancer marker, allergen, drug  
20 or other moiety having an ability to specifically bind with an antibody or antibody-like agent encoded by a multifunctional polypeptide of the present invention. For instance, the second nucleotide sequence can encode an anti-CEA antibody when the target ligand is CEA  
25 (carcinoembryonic antigen). The ligand can also be a fragment of DNA or other nucleic acids.

Nucleic acid probes specific for a portion of nucleotides that encode calf IAP can be used to detect nucleic acids specific to calf IAP for diagnostic purposes.  
30 Nucleic acid probes suitable for such purposes can be prepared from the cloned sequences or by synthesizing oligonucleotides that hybridize only with the homologous sequence under stringent conditions. The oligonucleotides can be synthesized by any appropriate method, such as by an  
35 automated DNA synthesizer.

The oligonucleotides can be used to detect DNA and mRNA or to isolate cDNA clones from libraries. The particular nucleotide sequences selected are chosen so as to correspond to the codons encoding a known amino acid sequence from the protein. Generally, an effective length of a probe is recognized in the art is about 14 to about 20 bases. Longer probes of about 25 to about 60 bases can also be used. A probe can be labelled, using labels and methods well known in the art, such as a radionucleotide or biotin, using standard procedures.

The purified recombinant calf IAP or its active fragments can be used for diagnostic purposes to determine the presence or concentration of a ligand in a sample. The sample can be a fluid or tissue specimen obtained, for example, from a patient suspected of being exposed to a particular antigen or DNA fragment. Those skilled in the art will recognize that any assay capable of using an enzyme-catalyzed system can be used in the detection methods of the present invention.

In the detection methods of the present invention:

(a) a sample is contacted with the recombinant calf IAP or an active fragment thereof attached to a reagent specifically reactive with the ligand to be detected;

(b) the sample is contacted with a detectable agent catalyzed by calf IAP; and

(c) the binding of the sample to the reagent is detected, where binding indicates the presence of the ligand in the sample.

The methods can also be used to determine the concentration of a ligand in the sample by relating the amount of binding to the concentration of the ligand. To determine the concentration, the amount of binding can be

compared to known concentrations of the ligand or to standardized measurements, such as slopes, determined from known concentrations of the ligand.

5 A variety of ligands can be detected by the present methods. The ligand can be, for example, a protein or polypeptide having antigenic properties or a nucleic acid, such as DNA or RNA.

10 Reagents reactive with such ligands can be antibodies or reactive fragments of such antibodies when the ligand is an antigen or antigen-like molecule. The reagent can also be a nucleotide probe that hybridizes or binds to a specific nucleic acid, such as DNA or RNA. Such probes can be oligonucleotides that are complementary to cDNA or genomic fragments of a ligand.

15 Procedures for attaching the enzymes to various reagents are well known in the art. Techniques for coupling enzymes to antibodies, for example, are described in Kennedy et al., Clin. Chim. Acta 70:1 (1976), incorporated herein by reference. Reagents useful for such  
20 coupling include, for example, glutaraldehyde, p-toluene diisocyanate, various carbodiimide reagents, p-benzoquinone m-periodate, N,N'-o-phenylenediamaleimide and the like. Alternatively, the multifunctional polypeptides of the present invention can be used.

25 Suitable substrates for the biochemical detection of ligands according to the methods of the present invention include, for example, p-nitrophenylphosphate.

The recombinant form of calf IAP is also useful for the development of calf IAP having greater heat  
30 stability. By site directed mutagenesis, it is possible to modify the nucleic acid sequence encoding for the

recombinant protein to obtain a heat stable calf IAP comparable to human placental IAP, which is known to be stable at about 65°C. Greater heat stability would allow the use of such a modified calf IAP in procedures requiring higher heating, such as Southern blotting, for example, which generally denatures many enzymes.

The following examples are intended to illustrate but not limit the invention.

#### EXAMPLE I

##### 10                    Libraries and Screening Procedures

Initially, a  $\lambda$ gt11 cDNA library prepared from adult bovine intestine (Clontech Laboratories, Palo Alto, CA) was screened using a mouse IAP cDNA fragment described in Manes et al., Genomics 8:541-554 (1990) as a probe. A 2.1 kb unprocessed cDNA fragment and a 1.1 kb processed cDNA fragment, both isolated from this library, were used to screen a genomic library prepared from adult cow liver in EMBL3 SP6/T7 (Clontech Laboratories, Palo Alto, CA). Radiolabelling of probes with  $^{32}\text{P}$  and identification and isolation of positive clones was done as described in Manes et al., supra, which is incorporated herein by reference. Large-scale phage DNA preparation was performed as described in Sambrook et al., supra, incorporated herein by reference.

25                    Initially, one positive cDNA clone was obtained upon screening the  $\lambda$ gt11 cDNA library with the mouse IAP cDNA fragment. Sequencing from the ends of the 2.1 kb cDNA fragment (R201) revealed an incomplete cDNA encoding exons VI through XI of an alkaline phosphatase gene as identified by sequence comparison to known AP genes. This cDNA fragment included all introns and revealed several STOP codons as well as two frameshifts in the putative coding region of the gene.



Although further sequence information of R201 suggested that it is possibly transcribed from a pseudogene, it was used as a probe for further screening of the  $\lambda$ gt11 library. Two additional cDNA clones were subsequently isolated and identified as transcripts of another alkaline phosphatase gene. Again, one fragment of 0.8 kb length (BB203) turned out to be reverse transcribed from an incomplete and unprocessed RNA, whereas the other one, a cDNA fragment of 1.1 kb length (BB204), was derived from a partial but processed mRNA, extending from the end of exon V through exon XI, lacking a putative polyadenylation site and a poly-A tail.

#### EXAMPLE II

##### Characterization of Genomic Clones and Sequence Analysis

Genomic DNA was isolated from adult cow liver and Southern blot analysis was performed using standard protocols as described in Sambrook et al., supra. Restriction enzymes were obtained from Gibco BRL, Boehringer Mannheim, and New England Biolabs. Twenty  $\mu$ g of genomic DNA were used per reaction. The blots were probed with the 2.1 kb unprocessed cDNA fragment, and washed under high stringency conditions (0.1 x SSC at 65°C).

Two bands in the genomic Southern were identified as fragments derived from the b.IAP gene. The only other non-human mammalian genome investigated extensively for tissues specific (TSAP) genes so far has been the murine genome, as reported in Manes et al., supra. Two murine TSAP genes, one termed embryonic AP (EAP), the other coding for IAP, and a pseudogene were cloned. In previous studies, it was shown that there are two TSAP genes expressed in the bovine genome according to Culp et al., Biochem. Biophys. Acta 831:330-334 (1985) and Besman & Coleman, supra. Similarly, two APs have been found expressed in the adult intestine of mice as reported in



Hahnel et al., Development 110:555-564 (1990). Expression of AP in rat intestine appears to be even more complex (Ellakim et al., Am. J. Physiol. 159, 1.1:G93-98 (1990)). Identification of the b.IAP gene was possible by comparison  
5 of its deduced amino acid sequence with N-terminal sequences reported for both TSAP isozymes.

Since further screening of the cDNA library revealed no additional positive clones, both R201 and BB204 were used to screen an EMBL3 SP6/T7 genomic library. Three  
10 positive clones were obtained and analyzed by Southern blotting. Subsequent sequencing of several fragments from two of the clones showed that one contained the entire coding region for the b.IAP gene as identified by comparison of deduced amino acid sequence with sequences  
15 previously determined in Culp et al., supra and Besman & Coleman, supra. A 5.4 kb sequence from overlapping Hind III and BamHI fragments of the clone containing the b.IAP gene are presented in Figure 1. The other clone contained sequences identical (except for a few basepair changes)  
20 with R201.

Genomic clones were characterized and sequences were determined as described in Manes et al., supra. Nucleic acid and protein sequences were assembled and analyzed using the MacVector sequence analysis program  
25 (IBI, New Haven, CT).

### EXAMPLE III

#### PCR Mutagenesis and Subcloning into pCDNA

A 23-mer primer ("MKNHE" (SEQ ID NO: 1):5'-GCTAGCCATGCAGGGGGCCTGCG-3' (SEQ ID NO: 2)) was used to  
30 amplify base pairs 1497-1913 of the b.IAP gene which had been subcloned as a Hind III/BamHI fragment into Bluescript-KS+ (Stratagene, San Diego, CA). MKNHE (SEQ ID NO: 1) had been designed to create a new Nhe I site by

altering the three 5' nucleotides of the primer sequence compared to the genomic sequence to allow the easy subcloning into different expression vectors. The universal SK primer was used as complementary reverse primer in the performed polymerase chain reaction (PCR). The plasmid was heat denatured, annealed to the primers and subjected to 30 cycles of PCR amplification in an Automatic Thermocycler (MJ Research, Piscataway, NJ). Times and temperatures were set as follows: annealing at 40°C for 30 seconds, extension for 3 minutes at 72°C and denaturing at 95°C for 30 seconds. The amplified fragment was directly subcloned into the "T-modified" EcoRV site of Bluescript as described in Marchuk et al., Nucl. Acids Res. 19:1154 (1990), incorporated herein by reference, in the orientation of b-galactosidase transcription.

#### EXAMPLE IV

##### Sequencing of the Amplified Fragment

The amplified fragment was sequenced using the universal T3 and T7 primers in the Sanger dideoxy chain termination procedure as described in Sanger et al., Proc. Natl. Acad. Sci. U.S.A. 74:5463-5467 (1977), which is incorporated herein by reference, to exclude the possibility of secondary mutations. The Hind III/BamHI fragment was used together with a 3.2 kb BamHI/SmaI fragment of the b.IAP gene for directional subcloning into a Hind III/EcoRV opened pCDNA 1 expression vector (Invitrogen, San Diego, CA).

#### EXAMPLE V

##### Recombinant Expression of b.IAP

The b.IAP gene subcloned into pCDNA 1 was transfected into Chinese hamster ovary (CHO) cells, ATCC No. CCL61, by means of Ca<sup>2+</sup> coprecipitation as described in Hummer and Millan, Biochem. J. 274:91-95 (1991), which is

incorporated herein by reference. The recombinant protein was extracted with butanol after incubating for 2 days.

The b.IAP gene presented in Figure 1 includes an open reading frame (ORF) of 2946 bp, containing 11 exons and 10 introns of very compact nature. Exon and intron borders were determined by comparison with BB204 and other known AP genes described in Manes et al., supra, Hernthorn et al., J. Biol. Chem. 263:12011-12019 (1988), Knoll et al., J. Biol. Chem. 263:12020-12027 (1988), and Millan & Manes, Proc. Natl. Acad. Sci. USA 85:3025-3028 (1988). A translation initiation codon ATG was identified by sequence comparison to known TSAP genes and is preceded by an in-frame STOP codon 48 bp upstream. The ORF, which is terminated by the STOP codon TAA, codes for a peptide of 533 amino acids in length. The mature protein of 514 amino acids with a calculated  $M_r$  of 64,400 Da is preceded by a hydrophobic signal peptide as is the case for all known APs.

The predicted amino acid sequence of the b.IAP protein is highly homologous to other known IAPs as shown in Figure 3. As shown in Figure 3 there is identity in those parts corresponding to the partial amino acid sequences previously determined for b.IAP (Culp et al., supra; Besman and Coleman, supra). Besman & Coleman determined N-terminal amino acid sequences for two differentially expressed AP isozymes. The 16 N-terminal amino acids determined for the isozyme found only in newborn calves differ in three or four residues from the N-terminus of the enzyme exclusively expressed in adults.

30

#### EXAMPLE VI

##### Reverse Transcriptase-PCR

In order to construct a full length cDNA, reverse transcriptase-PCR (RT-PCR) was performed as follows: total

RNA from a stable transfected CHO-cell clone (M2) was isolated by acid guanidium thiocyanate-phenol-chloroform extraction as described in Chomczynski & Sacchi, Anal. Biochem. 162:156-159 (1987), incorporated herein by reference. The reverse transcriptase reaction was conducted according to the protocol of the manufacturer (Promega, Wisconsin) using 10 µg of RNA.

The reaction mixture was extracted with phenol-chloroform, precipitated with ethanol and resuspended in Taq polymerase buffer. The subsequent PCR was performed over 35 cycles of amplification following an initial denaturation at 94°C for 5 minutes, annealing at 55°C for 30 seconds and extension at 72°C for 5 minutes. The Taq Polymerase was added to the reaction mixture after denaturation only. The subsequent PCR settings were: denaturation at 94°C for 45 seconds, annealing at 55°C for 1 minute and extension at 72°C for 4 minutes. The primers used for this reaction were MKNHE (SEQ ID NO: 1) and sequencing primer UP6: TCGGCCGCCTGAAGGAGC (SEQ ID NO: 3) (see Figure 2).

The sequencing strategy as well as a restriction map and the genomic structure of the b.IAP gene are shown in Figure 2. The strategies for subcloning the coding region of the gene into an expression vector using PCR and for construction of a full length cDNA by means of RT-PCR are indicated in Figure 2. A single fragment of approximately 830 bp length had been obtained from RT-PCR as could be expected from the genomic sequence.

#### EXAMPLE VIII

##### 30      Characterization of Recombinant Calf IAP

The sequence for the calf intestinal AP gene was determined as described above. A full length cDNA was constructed using a partial cDNA clone (BB204) and a

fragment obtained by RT-PCR.

A cDNA fragment clone (R201) and a corresponding genomic clone were obtained, which resemble properties of a putative pseudogene. Both clones contain STOP codons within the coding region and several frameshifts. Bands corresponding to the putative pseudogene could only be identified upon hybridizing with a mouse TNAP cDNA which gave a distinct pattern. This result suggests that the bands correspond to TSAP genes only, and that the pseudogene is more related to TNAP. In contrast, the murine pseudogene has been found to resemble more homology to the mouse EAP gene (Manes et al., supra).

The sequence and genomic structure of the b.IAP gene show high homology to all known TSAP genes. The smallest exon, exon VII, is only 73 bp long while the longest exon, exon XI, is approximately 1.1 kb long. The exact length of exon 11 cannot be determined since no cDNA with a poly-A tail had been isolated. The estimate given is based on the identification of a putative polyadenylation site AATAAA (bp 5183-5188) in the 3' non-coding region of the gene (underlined in Figure 1). The introns are among the smallest introns reported (Hawkins, Nucl. Acids Res. 16:9893-9908 (1988)) as was found in the case of other TSAP genes as well (Manes et al., supra; Hernthorn et al., supra; Knoll et al., supra; Millan and Manes, supra). The largest one, splitting exon V and exon VI, is only 257 bp long. All exon-intron junctions conform to the GT-AG rule (Breathnach et al., Proc. Natl. Acad. Sci. USA 75:4853-4857 (1978)) and also conform well to the consensus sequences (C/A)AG/GT(A/G)AGT (SEQ ID NO: 4) and (T/C)<sub>n</sub>N(C/T)AG/G (SEQ ID NO: 5) for donor and acceptor sites, respectively (Mount, Nucl. Acids Res. 10:459-473 (1982)).

Interestingly, the entire coding region of exon



XI shows a high G/C content of over 60 to 80% compared to a rather equal ratio of G/C to A/T throughout the whole structural gene. Other regions of biased GC content were found at bp 270 to bp 490 with a high A/T content and in a region preceding the poly adenylation site, which again shows a high G/C content.

A putative TATA-box has been identified in the 1.5 kb of sequence preceding the coding region (bp 1395-1400, underlined in Figure 1). It shows the same variant ATTTAA sequence embedded in a conserved region of 25 bp as was previously reported for the mouse TSAP genes (Manes et al., supra) and two human TSAP genes (Millan, Nucl. Acids Res. 15:10599 (1987); Millan and Manes, supra)).

The sequence GGGAGGG has been shown to be part of the putative mouse TSAP promoters (Manes et al., supra) as well as of two human TSAP promoters (Millan, (1987), supra; Millan and Manes, supra). This sequence is also present in the putative promoter region of the b.IAP gene.

The sequence CACCC or its complementary reverse is repeated 6 times in the region of bp 1182-1341, 24 times in the entire structural gene and 31 times throughout the whole sequence shown here. However, only one less conserved CACCC box (Myers et al., Science 232:613-618 (1986)) was identified.

Since it was shown for dog IAP that the enzyme can be induced by cortico steroid hormone (Sanecki et al., Am. J. Vet. Res. 51, 12:1964-1968 (1990)), hormone responsive elements in the genomic sequence of b.IAP were identified. Palindromic and direct repeats, known to be binding sites for dimeric nuclear factors as described in O'Malley, Mol. Endocrinol. 5:94-99 (1990), were identified in the 1.5 kb upstream of the initiation codon. A long, imperfect palindromic repeat (CACACCTCCTGCCCAG-N,-



CTGGTGAGGAGCTGAG) (SEQ ID NO: 6) extends from bp 899 to bp 937. A direct repeat of the sequence GGGCAGG spaced by three nucleotides starts at bp 1311.

Several regions of high homology to mouse (Manes et al., supra) and human (Millan, (1987), supra) IAP genes have been identified in the putative promoter region. However, one stretch of 10 bp (AGCCACACCC) (SEQ ID NO: 7) was found to be identical with a sequence in the same region upstream of the TATA box of the human  $\beta$ -globin gene (Myers et al., supra).

Another region of interest precedes the putative poly adenylation site at bp 5016. The sequence ACAGAGAGGAGA (SEQ ID NO: 8) is imperfectly repeated, spaced by an inverted repeat overlapping the last adenine nucleotide (ACAG-T-GACA). The presented 1.5 kb of the presumed promoter of the b.IAP gene contain several additional putative regulatory elements. A short stretch of 14 alternating thymines and guanines, intercepted by one adenine was found at position 601 of the sequence. Interestingly, this sequence is identical to a part of a slightly longer stretch with the same characteristics beginning at bp 2713 within the intron splitting exon V and VI. Another stretch of 36 alternating pyridines and purines is found at position 732 being mainly composed of cytosin and adenine nucleotides. Identical structures are reported for the human germ cell AP gene (Millan and Manes, supra) and are thought to form Z-DNA structures, which may play a role in the regulation of gene expression (Nordheim and Rich, Nature (London) 303:674-678 (1983)).

As shown in Figure 3, the deduced amino acid sequence of b.IAP is highly homologous to all known IAPs. Identical residues and conservative amino acid substitutions are found within structurally important regions, as is the case for the other TSAPs as well,

whereas variability is almost exclusively found at the C-terminus and in the highly variable loops (Millan, (1988), supra).

Asp<sup>487</sup> of b.IAP resides within a conserved sequence  
5 of 4 amino acids in the same region of the human intestinal  
gene (indicated in Figure 3) as well as of human PLAP  
(Millan, J. Biol. Chem. 261:3112-3115 (1986)). This  
residue was shown for PLAP to be the attachment site of a  
phosphatidyl-inositol membrane anchor (Micanovic et al.,  
10 Proc. Natl. Acad. Sci. USA 87:157-161 (1990)). Evidence  
has been presented previously that b.IAP is also anchored  
to the plasma membrane in such a fashion. There appears to  
be a spatial regulated release of IAP into the lumen  
without cleavage of the anchor in a variety of species  
15 (Hoffmann-Blume et al., Eur. J. Biochem. 199:305-312  
(1991)).

#### EXAMPLE IX

##### Comparison of Purified and Recombinant Forms of Calf IAP

Values for  $K_m$  and  $K_i$  for L-Phe were determined for  
20 the recombinant enzyme as well as for purified protein from  
calf intestine as described in Hummer and Millan, supra,  
and Wilkinson, Biochem. J. 8:324-332 (1961), incorporated  
herein by reference. Both the purified b.IAP from natural  
sources and the recombinant b.IAP show identical values for  
25  $K_m$  (within standard deviations), and only slightly different  
values of  $K_i$ .  $K_m$  was determined as  $0.77 \pm 0.12$  for the  
recombinant enzyme and as  $0.86 \pm 0.17$  for the purified  
natural enzyme.  $K_i$  for L-Phe were found to be  $15.2 \pm 1.8$   
and  $11.2 \pm 1.0$  for the recombinant and purified enzymes,  
30 respectively. Thus, the results of these findings indicate  
that the natural and recombinant forms of calf IAP have  
comparable properties and activities.

Two possible glycosylation sites appear to be

conserved between the human and the bovine IAP. Three other possible sites within other IAP sequences were not found in the b.IAP. The high degree of heterologous glycosylation of the purified enzyme was demonstrated by isoelectric focusing (IEF). IEF was performed using the Resolve-ALP system (Isolab, Akron, OH) as described in Griffiths & Black, Clinn. Chem. 33:2171-2177 (1987). Samples of recombinant and purified enzyme were run either treated with neuraminidase or untreated to compare the amount of glycosylation.

A smeary band was obtained upon IEF of untreated purified enzyme in contrast to a more distinct band for the recombinant b.IAP protein. After treatment with neuraminidase, both bands dissolve into several sharp bands, in which the purified enzyme showed considerably more diversity than the recombinant enzyme.

#### EXAMPLE X

##### Heat Inactivation of Calf IAP

The heat stabilities of purified calf IAP and recombinant calf IAP were determined at 56°C. First, the enzyme samples were diluted in 1 ml of DEA buffer containing 1 M DEA diethanolamine (pH 9.8) containing 0.5 mM MgCl<sub>2</sub> and 20 µM ZnCl<sub>2</sub>. The solution was heated at 56°C for the fixed time intervals indicated in Table I. Fifty µl of the enzyme solution were removed and pipetted into a microtiter well and stored on ice until the end of the longest incubation period. At the end of the experiment, the residual activity was measured by the addition of 200 µl of DEA buffer containing p-nitrophenylphosphate (10 mM) in DEA buffer. For comparison, a sample of recombinant enzyme was pretreated with 0.2 units/ml of neuraminidase for 16 hours at room temperature, followed by the same heat inactivation treatment. The results of the heat inactivation studies are shown in Figure 4.

TABLE IHeat Inactivation of Intestinal AP

		<u>Time (minutes)</u>					
		<u>0'</u>	<u>6'</u>	<u>12'</u>	<u>18'</u>	<u>24'</u>	<u>30'</u>
5		<u>Residual activity (%)</u>					
	Calf IAP (intestinal extract)	100	87	65.6	48.7	36	23.4
	Recombinant IAP	100	80.6	59.5	39.6	28.5	18.5
10	Recombinant IAP upon Neuriminidase	100	80.8	55.9	38.1	27.1	20.3

The foregoing description of the invention is  
15 exemplary for purposes of illustration and explanation. It  
should be understood that various modifications can be made  
without departing from the spirit and scope of the  
invention. Accordingly, the following claims are intended  
to be interpreted to embrace all such modifications.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

5 (A) NAME: La Jolla Cancer Research Foundation  
(B) STREET: 10901 North Torrey Pines Road  
(C) CITY: La Jolla  
(D) STATE: California  
(E) COUNTRY: USA  
10 (F) POSTAL CODE (ZIP): 92037  
(G) TELEPHONE: (619) 455-6480  
(H) TELEFAX: (619) 455-0181

(ii) TITLE OF INVENTION: RECOMBINANT CALF INTESTINAL ALKALINE  
PHOSPHATASE

15 (iii) NUMBER OF SEQUENCES: 13

## (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
20 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version 1.25 (EPO)

## (vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/849,219  
(B) FILING DATE: 10-MAR-1992

## (2) INFORMATION FOR SEQ ID NO:1:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

30 Met Lys Asn His Glu  
1 5

## (2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:  
35 (A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCTAGCCATG CAGGGGGCCT GCG

23

## 40 (2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
45 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TCGGCCGCCT GAAGGAGC

18

## (2) INFORMATION FOR SEQ ID NO:4:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 10 (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION: complement (1)  
(D) OTHER INFORMATION: /note= "N=C OR A"

- 15 (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION: complement (2)  
(D) OTHER INFORMATION: /note= "N=AG OR GT"

- 20 (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION: complement (3)  
(D) OTHER INFORMATION: /note= "N=A OR G"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

NNNAGT

6

## (2) INFORMATION FOR SEQ ID NO:5:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 30 (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION: complement (1)  
(D) OTHER INFORMATION: /note= "Y=T OR C"

- 35 (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION: complement (3)  
(D) OTHER INFORMATION: /note= "Y=C OR T"

- 40 (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION: complement (4)  
(D) OTHER INFORMATION: /note= "Y=AG OR G"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

YNY Y

4



25

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CACACCTCCT GCCCAGNNNN NNNCTGGTGA GGAGCTGAG

39

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGCCACACCC

10

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ACAGAGAGGA GA

12

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5399 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

30

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(1501..1567, 1647..1763, 1878..1993, 2179  
..2353, 2433..2605, 2864..2998, 3084..3156, 3257  
..3391, 3475..3666, 3879..3995, 4101..4402)

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAGCTTTCAC CTTCTCTGAA AACAGAGAGA CAGTCCTCAG CCCAGTCCT CACCCTTCCT

60

ACCTCCCTGC CTGATGCCCA GGCAATCATC TGGTGGCGTG TCACCTCCCT CTGTCCCATG

120

40 AGTTCCACTA GATGTGGCCC TCAAGAAAAA GGGCTTCCCT GTTGGCTCAG CTGGTAAAGA

180

	ATCCTCCAGC AATGTAGGAG ACCTGGGTTC GATCCCTGGG TTGGGAGGAT ACCCTGGAGA	240
	AGGGAATGGC TACCCACTCC AGTATTCTTG CCTGGATAAT CCCATGGACA GAGGAGTCTG	300
	GCAGGCTGCA GACCATAAGG TAGAAAGAGT CAGACATGAC TGAGCAACTA AGCACAATAT	360
	TCCACTGGAT ATATCATACT TTGTTTCATCC ATTTGTCTGC TGTGGATGGT TGAGTGGCTT	420
5	GTGCCTCTTG GCTACTGTGA GTAATGCTAC TAAAATGTGA GTGTGCAAAT ACCTCTTATA	480
	GATCTTGATT TCAATTATTG GGGATACACA CCCAGAAGGC GGATTGTTGG ATGTGAGAAT	540
	GCCTTTTTGA ACCCCAACCT GGGGTTACTG AAACCCTAGC TCCTTATCAG AAGCTGTTCC	600
	TGTGAGTGTG TGTGGCCTGT GGAGAGAAGA GACTCACCTC TGCCTTCCAT TTACCTCTCC	660
	AATGGAGCAG AGGTTGCAAA CTTCAGTTAA TGGGCACTGG GCCCACGCCT GTCGACCCGT	720
10	TACAGGCACC TTACACACAC ACACACACAC ACACACACAC ACAAACAGCA CTGCAGACCC	780
	AGCTCTTCAG TAACTGAAGA CACAGACAAG GCCCCGCTC TGCTGTCACC TCCAGTCCCA	840
	TCCTTCTCCA CAGCAGAAGC TGGGCCCAGG CTCCCATGTG CCCCCACTAG CCCAGTGCCC	900
	ACACCTCCTG CCCAGGTCAA GTCTGGTGAG GAGCTGAGCA GGGGGCAGGG CAGACAGGCC	960
	TCCCCGTGGA TCTCTGTCTC AGGGCGCCAG GGAACCTAACC CAGGCCCCTG GCCAGGCTGT	1020
15	GTCCCTAAGC ACTGGGAACC AAACCAGGCC AAGGCTGAGT CTCAGAAAAC ACTGAACACG	1080
	TGAAGGAAGG AGAGATGGTT CTCCCACAGG ACTTGGTGAG CAGAGGGCTG GGAGGAGCCT	1140
	CAGTCAGGAC CTTGAAAACG TTCCTCAGGC CTAGACATCT GCACCCTAAT CCCCACCCCA	1200
	CCCTGAGGAG ACAGCTGGGA CCATCCTGGG AGGGAGGGAC CTGAATCCTC AGGACCCCTA	1260
	CTGCTAAGCC ACACCCACCA CATGCCCCTG GCAACAGGGC TCAAAGTCAT AGGGCAGGTG	1320
20	AGGGGCAGGG TGTGGCCACC CGGGGAACCT GGGATGGACA AGGAGACTTT AATAGCAGGG	1380
	ACAAAGTCTA TCTAGATTTA AGCCCAGCAG GCCAAGCTGC AGCCGGTCCC TGGTGTCCCA	1440
	GCCTTGCCCT GAGACCCGGC CTCCCCAGGT CCCATCCTGA CCCTCTGCCA TCACACAGCC	1500
	ATG CAG GGG GCC TGC GTG CTG CTG CTG CTG GGC CTG CAT CTA CAG CTC	1548
25	Met Gln Gly Ala Cys Val Leu Leu Leu Leu Gly Leu His Leu Gln Leu	
	1 5 10 15	
	TCC CTA GGC CTC GTC CCA G GTAATCAGGC GGCTCCCAGC AGCCCCTACT	1597
	Ser Leu Gly Leu Val Pro	
	20	
30	CACAGGGGCG GCTCTAGGCT GACCTGACCA ACACTCTCCC CTTGGGCAG TT GAG	1651
	Val Glu	
	GAG GAA GAC CCC GCC TTC TGG AAC CGC CAG GCA GCC CAG GCC CTC GAT	1699
	Glu Glu Asp Pro Ala Phe Trp Asn Arg Gln Ala Ala Gln Ala Leu Asp	
	25 30 35 40	
35	GTG GCT AAG AAG CTG CAG CCC ATC CAG ACA GCC GCC AAG AAT GTC ATC	1747
	Val Ala Lys Lys Leu Gln Pro Ile Gln Thr Ala Ala Lys Asn Val Ile	
	45 50 55	
	CTC TTC TTG GGG GAT G GTGAGTACAT GAGGCCAGCC CACCCCCTGT	1793
40	Leu Phe Leu Gly Asp	
	60	

	CCCCTGACAG GCCTGGAACC CTGTGATGCC GGCTGACCCA GGTGGCCCC AGAAACTCGG	1853
	ACCTGAGACA CTGTGTACCT TCAG GG ATG GGG GTG CCT ACG GTG ACA GCC Gly Met Gly Val Pro Thr Val Thr Ala	1903
		65 70
5	ACT CGG ATC CTA AAG GGG CAG ATG AAT GGC AAA CTG GGA CCT GAG ACA Thr Arg Ile Leu Lys Gly Gln Met Asn Gly Lys Leu Gly Pro Glu Thr	1951
		75 80 85
10	CCC CTG GCC ATG GAC CAG TTC CCA TAC GTG GCT CTG TCC AAG Pro Leu Ala Met Asp Gln Phe Pro Tyr Val Ala Leu Ser Lys	1993
		90 95 100
	GTAAGGCCAA GTGGCCTCAG GGTGGTCTAC ACCAGAGGGG TGGGTGTGGG CCTAGGGAGC	2053
	AGGGTAGGAG GGAAACCCAG GAGGGCTAGG GGCTGAGATA GGGGCTGGGG GCTGTGAGGA	2113
	TGGGCCCAGG GCTGGGTCAG GAGCTGGGTG TCTACCCAGC AGAGCGTAAG GCATCTCTGT	2173
15	CCCAG ACA TAC AAC GTG GAC AGA CAG GTG CCA GAC AGC GCA GGC ACT Thr Tyr Asn Val Asp Arg Gln Val Pro Asp Ser Ala Gly Thr	2220
		105 110
	GCC ACT GCC TAC CTG TGT GGG GTC AAG GGC AAC TAC AGA ACC ATT GGT Ala Thr Ala Tyr Leu Cys Gly Val Lys Gly Asn Tyr Arg Thr Ile Gly	2268
		115 120 125 130
20	GTA AGT GCA GCC GCC CGC TAC AAC CAG TGC AAA ACG ACA CGT GGG AAT Val Ser Ala Ala Ala Arg Tyr Asn Gln Cys Lys Thr Thr Arg Gly Asn	2316
		135 140 145
25	GAG GTC ACG TCT GTG ATG AAC CGG GCC AAG AAA GCA G GTGGGCTTGG Glu Val Thr Ser Val Met Asn Arg Ala Lys Lys Ala	2363
		150 155
	GCGTCAGCTT CCTGGGCAGG GACGGGCTCA GAGACCTCAG TGGCCCACCG TGACCTCTGC	2423
	CACCCTCAG GG AAG TCC GTG GGA GTG GTG ACC ACC ACC AGG GTG CAG Gly Lys Ser Val Gly Val Val Thr Thr Thr Arg Val Gln	2470
		160 165 170
30	CAT GCC TCC CCA GCC GGG GCC TAC GCG CAC ACG GTG AAC CGA AAC TGG His Ala Ser Pro Ala Gly Ala Tyr Ala His Thr Val Asn Arg Asn Trp	2518
		175 180 185
35	TAC TCA GAC GCC GAC CTG CCT GCT GAT GCA CAG ATG AAT GGC TGC CAG Tyr Ser Asp Ala Asp Leu Pro Ala Asp Ala Gln Met Asn Gly Cys Gln	2566
		190 195 200
	GAC ATC GCC GCA CAG CTG GTC AAC AAC ATG GAT ATT GAC GTGCGACATG Asp Ile Ala Ala Gln Leu Val Asn Asn Met Asp Ile Asp	2615
		205 210 215
	TTGGGCACAG GGCGGGGCTG GGCACAGGTG GTGGGGCACA CTCGCAACAC AGTCGTAGGT	2675
40	AACCTCCAGC CTGCGGTGTT TCAGGGTTTT CATGGGTTTG TGTGTGTGTG TATGTGTGGT	2735
	GGGGTGGCAC CATGTAGGAG GTGGGGACAG GCCTTTCCCA CAGACCTGGT GGGGGAGGTA	2795
	GGGGCTGTGT GAGAGGAGTA AAGGGCCAGC CAGGCCCTA ACCCACCTGC CTAACCTCTCT	2855
45	GGCTCCAG GTG ATC CTG GGT GGA GGC CGA AAA TAC ATG TTT CCT GTG GGG Val Ile Leu Gly Gly Arg Lys Tyr Met Phe Pro Val Gly	2905
		220 225 230

	ACC CCA GAC CCT GAA TAC CCA GAT GAT GCC AGT GTG AAT GGA GTC CGG	2953
	Thr Pro Asp Pro Glu Tyr Pro Asp Asp Ala Ser Val Asn Gly Val Arg	
	235 240 245	
5	AAG CGA AAG CAG AAC CTG GTG CAG GCA TGG CAG GCC AAG CAC CAG	2998
	Lys Arg Lys Gln Asn Leu Val Gln Ala Trp Gln Ala Lys His Gln	
	250 255 260	
	GTAATGGGGG CTCACGGATG TGGGGGTACA GTGGGGCTGG GCCTGGGGTG TCGGCTATGG	3058
	CTGAGGCCTG GTTCTGCCCT CCCAG GGA GCC CAG TAT GTG TGG AAC CGC ACT	3110
10	Gly Ala Gln Tyr Val Trp Asn Arg Thr	
	265 270	
	GCG CTC CTT CAG GCG GCC GAT GAC TCC AGT GTA ACA CAC CTC ATG G	3156
	Ala Leu Leu Gln Ala Ala Asp Asp Ser Ser Val Thr His Leu Met	
	275 280 285	
	GTAACGACTC CACCCACCCT CACTGTCCTC CCCAGGAATG GGTGCCATGG GCCACCCCTG	3216
15	TCCTCAGCTT GAGGGTCACC ACTGCTCCCC TTTCCACAG GC CTC TTT GAG CCG	3270
	Gly Leu Phe Glu Pro	
	290	
	GCA GAC ATG AAG TAT AAT GTT CAG CAA GAC CAC ACC AAG GAC CCG ACC	3318
20	Ala Asp Met Lys Tyr Asn Val Gln Gln Asp His Thr Lys Asp Pro Thr	
	295 300 305	
	CTG CAG GAA ATG ACA GAG GTG GCC CTG CGA GTC GTA AGC AGG AAC CCC	3366
	Leu Gln Glu Met Thr Glu Val Ala Leu Arg Val Val Ser Arg Asn Pro	
	310 315 320	
	AGG GGC TTC TAC CTC TTT GTG GAG G GTGAGTGGCA GCCCCTTGGT	3411
25	Arg Gly Phe Tyr Leu Phe Val Glu	
	325 330	
	GAACAGAGGT GTGATGAGGG CCATCAGGGT GGGTTTGGTA TCTTATATGT GACTTATCTG	3471
	CAG GA GGC CGC ATT GAC CAC GGT CAC CAT GAT GAC AAA GCT TAT ATG	3518
30	Gly Gly Arg Ile Asp His Gly His His Asp Asp Lys Ala Tyr Met	
	335 340 345	
	GCA CTG ACC GAG GCG GGT ATG TTT GAC AAT GCC ATC GCC AAG GCT AAT	3566
	Ala Leu Thr Glu Ala Gly Met Phe Asp Asn Ala Ile Ala Lys Ala Asn	
	350 355 360	
	GAG CTC ACT AGC GAA CTG GAC ACG CTG ATC CTT GTC ACT GCA GAC CAC	3614
35	Glu Leu Thr Ser Glu Leu Asp Thr Leu Ile Leu Val Thr Ala Asp His	
	365 370 375	
	TCT CAT GTC TTC TCT TTT GGT GGC TAT ACA CTG CGT GGG ACC TCC ATT	3662
	Ser His Val Phe Ser Phe Gly Gly Tyr Thr Leu Arg Gly Thr Ser Ile	
	380 385 390	
40	TTT G GTAAGCCCAG GGAGAGTGGC AGGTCGTTGC CCCTAAGTTA CGAGGCACAA	3716
	Phe	
	CTCGTCTGAG CCAGTTCCTC TATCTGTCTA GTGGGGTAGT ACAGCACACT GCCTGCTACG	3776
	CTCTGGTGAG GATTGTCACT GACAGACAGA CTGGCCATGG CTCTGCACAC AGGGGAGCAC	3836
45	AAGCTAGGTC AGTGTGATCA CGGGGTCCCC TCTTCCCTGA AG GT CTG GCC CCC	3889
	Gly Leu Ala Pro	
	395	

	AGC AAG GCC TTA GAC AGC AAG TCC TAC ACC TCC ATC CTC TAT GGC AAT	3937
	Ser Lys Ala Leu Asp Ser Lys Ser Tyr Thr Ser Ile Leu Tyr Gly Asn	
	400 405 410	
5	GGC CCA GGC TAT GCG CTT GGC GGG GGC TCG AGG CCC GAT GTT AAT GAC	3985
	Gly Pro Gly Tyr Ala Leu Gly Gly Gly Ser Arg Pro Asp Val Asn Asp	
	415 420 425 430	
	AGC ACA AGC G GTAAGTGTAG TAGGTGGGGC GCTGGGAGGT GGGGACCCTG	4035
	Ser Thr Ser	
10	GCCAGAAATT GTGGGGAGGG GAAGGCTGCC TCCCTTGTC CATTAACTTC CCTTCTTCTG	4095
	GCCAG AG GAC CCC TCG TAC CAG CAG CAG GCG GCC GTG CCC CAG GCT	4141
	Glu Asp Pro Ser Tyr Gln Gln Gln Ala Ala Val Pro Gln Ala	
	435 440 445	
15	AGC GAG ACC CAC GGG GGC GAG GAC GTG GCG GTG TTC GCG CGC GGC CCG	4189
	Ser Glu Thr His Gly Gly Glu Asp Val Ala Val Phe Ala Arg Gly Pro	
	450 455 460	
	CAG GCG CAC CTG GTG CAC GGC GTC GAG GAG GAG ACC TTC GTG GCG CAC	4237
	Gln Ala His Leu Val His Gly Val Glu Glu Glu Thr Phe Val Ala His	
	465 470 475	
20	ATC ATG GCC TTT GCG GGC TGC GTG GAG CCC TAC ACC GAC TGC AAT CTG	4285
	Ile Met Ala Phe Ala Gly Cys Val Glu Pro Tyr Thr Asp Cys Asn Leu	
	480 485 490 495	
25	CCA GCC CCC ACC ACC GCC ACC AGC ATC CCC GAC GCC GCG CAC CTG GCG	4333
	Pro Ala Pro Thr Thr Ala Thr Ser Ile Pro Asp Ala Ala His Leu Ala	
	500 505 510	
	GCC AGC CCG CCT CCA CTG GCG CTG CTG GCT GGG GCG ATG CTG CTG CTG	4381
	Ala Ser Pro Pro Pro Leu Ala Leu Leu Ala Gly Ala Met Leu Leu Leu	
	515 520 525	
30	CTG GCG CCC ACC TTG TAC TAACCCCCAC CAGTTCCAGG TCTCGGGATT	4429
	Leu Ala Pro Thr Leu Tyr	
	530	
	TCCCGCTCTC CTGCCCAAAA CCTCCCAGCT CAGGCCCTAC CGGAGCTACC ACCTCAGAGT	4489
	CCCCACCCCG AAGTGCTATC CTAGCTGCCA CTCCTGCAGA CCCGACCCGG CCCCACCACC	4549
	AGAGTTTCAC CTCCCAGCAG TGATTCACAT TCCAGCATTG AAGGAGCCTC AGCTAACAGC	4609
35	CCTTCAAGGC CCAGCCTATA CCGGAGGCTG AGGCTCTGAT TTCCCTGTGA CACGCGTAGA	4669
	CCTACTGCCC GACCCCAACT TCGGTGGCTT GGGATTTTGT GTTCTGCCAC CCTGAACCTC	4729
	AGTAAGGGGG CTCGGACCAT CCAGACTGCC CCTACTGCCC ACAGCCCACC TGAGGACAAA	4789
	GCTGGCACGG TCCCAGGGGT CCCAGGCCCG GCTGGAACCC ACACCTTGCC TTCAGCGACC	4849
	TGGA CTCTGG GTTCGGAGAG TGGCTTCGGG AGGCGTGGTT TCCGATGGGC GTGCTCTGGA	4909
40	ACGTGCTCGC CTGAACCAAC CTGTGTACAC TGGCCAGGAA TCACGGCCAC CAGAGCTCGG	4969
	ACCTGACAGA GCCCTCAGCA GCCCCTCCTA GACCAACGTA CCCATTACAG AGAGGAGACA	5029
	GTGACACAGA GGAGAGGAGA CTTGTCCCAG GTCCCTCAGC TGCTGTGAGG GCGGCCCTGG	5089
	TGCCCCCTTC AGGCTGGGCA TCCCAGTAGC AGCAGGGGAC CCGGGGGTGG GGACACAGGC	5149
	CCCGCCCTCC CTGGGAGGCA GGAAGCAGCT CTCAAATAAA CTGTTCTAAG TATGATACAG	5209



GAGTGATACA TGTGTGAAGA GAAGCCCTTA GGTGGGGGCA CAGAGTGTCT GGGTGAGGGG 5269  
 GGTCAGGGTC ACATCAGGAG GTTAGGGAGG GGTTGATGAA GGGCTGACGT TGAGCAAAGA 5329  
 CCAAAGGCAA CTCAGAAGGA CAGTGGTGCA GGACTGGGTG TGGTCAGCAG GGGGACTGGT 5389  
 TGGGGGATCC 5399

5 (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 533 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Gln Gly Ala Cys Val Leu Leu Leu Leu Gly Leu His Leu Gln Leu  
 1 5 10 15

Ser Leu Gly Leu Val Pro Val Glu Glu Glu Asp Pro Ala Phe Trp Asn  
 15 20 25 30

Arg Gln Ala Ala Gln Ala Leu Asp Val Ala Lys Lys Leu Gln Pro Ile  
 35 40 45

Gln Thr Ala Ala Lys Asn Val Ile Leu Phe Leu Gly Asp Gly Met Gly  
 50 55 60

Val Pro Thr Val Thr Ala Thr Arg Ile Leu Lys Gly Gln Met Asn Gly  
 65 70 75 80

Lys Leu Gly Pro Glu Thr Pro Leu Ala Met Asp Gln Phe Pro Tyr Val  
 85 90 95

Ala Leu Ser Lys Thr Tyr Asn Val Asp Arg Gln Val Pro Asp Ser Ala  
 100 105 110

Gly Thr Ala Thr Ala Tyr Leu Cys Gly Val Lys Gly Asn Tyr Arg Thr  
 115 120 125

Ile Gly Val Ser Ala Ala Ala Arg Tyr Asn Gln Cys Lys Thr Thr Arg  
 130 135 140

Gly Asn Glu Val Thr Ser Val Met Asn Arg Ala Lys Lys Ala Gly Lys  
 145 150 155 160



31

	Ser Val Gly Val Val Thr Thr Thr Arg Val Gln His Ala Ser Pro Ala	165	170	175
5	Gly Ala Tyr Ala His Thr Val Asn Arg Asn Trp Tyr Ser Asp Ala Asp	180	185	190
	Leu Pro Ala Asp Ala Gln Met Asn Gly Cys Gln Asp Ile Ala Ala Gln	195	200	205
10	Leu Val Asn Asn Met Asp Ile Asp Val Ile Leu Gly Gly Gly Arg Lys	210	215	220
	Tyr Met Phe Pro Val Gly Thr Pro Asp Pro Glu Tyr Pro Asp Asp Ala	225	230	235
15	Ser Val Asn Gly Val Arg Lys Arg Lys Gln Asn Leu Val Gln Ala Trp	245	250	255
	Gln Ala Lys His Gln Gly Ala Gln Tyr Val Trp Asn Arg Thr Ala Leu	260	265	270
20	Leu Gln Ala Ala Asp Asp Ser Ser Val Thr His Leu Met Gly Leu Phe	275	280	285
	Glu Pro Ala Asp Met Lys Tyr Asn Val Gln Gln Asp His Thr Lys Asp	290	295	300
25	Pro Thr Leu Gln Glu Met Thr Glu Val Ala Leu Arg Val Val Ser Arg	305	310	315
30	Asn Pro Arg Gly Phe Tyr Leu Phe Val Glu Gly Gly Arg Ile Asp His	325	330	335
	Gly His His Asp Asp Lys Ala Tyr Met Ala Leu Thr Glu Ala Gly Met	340	345	350
35	Phe Asp Asn Ala Ile Ala Lys Ala Asn Glu Leu Thr Ser Glu Leu Asp	355	360	365
40	Thr Leu Ile Leu Val Thr Ala Asp His Ser His Val Phe Ser Phe Gly	370	375	380

32

Gly Tyr Thr Leu Arg Gly Thr Ser Ile Phe Gly Leu Ala Pro Ser Lys  
385 390 395 400

Ala Leu Asp Ser Lys Ser Tyr Thr Ser Ile Leu Tyr Gly Asn Gly Pro  
405 410 415

5

Gly Tyr Ala Leu Gly Gly Gly Ser Arg Pro Asp Val Asn Asp Ser Thr  
420 425 430

10

Ser Glu Asp Pro Ser Tyr Gln Gln Gln Ala Ala Val Pro Gln Ala Ser  
435 440 445

Glu Thr His Gly Gly Glu Asp Val Ala Val Phe Ala Arg Gly Pro Gln  
450 455 460

15

Ala His Leu Val His Gly Val Glu Glu Glu Thr Phe Val Ala His Ile  
465 470 475 480

Met Ala Phe Ala Gly Cys Val Glu Pro Tyr Thr Asp Cys Asn Leu Pro  
485 490 495

20

Ala Pro Thr Thr Ala Thr Ser Ile Pro Asp Ala Ala His Leu Ala Ala  
500 505 510

Ser Pro Pro Pro Leu Ala Leu Leu Ala Gly Ala Met Leu Leu Leu Leu  
515 520 525

25

Ala Pro Thr Leu Tyr  
530

## (2) INFORMATION FOR SEQ ID NO:11:

30

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 540 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

35

Met Gln Gly Asp Trp Val Leu Leu Leu Leu Leu Gly Leu Arg Ile His  
1 5 10 15

Leu Ser Phe Gly Val Ile Pro Val Glu Glu Glu Asn Pro Val Phe Trp  
20 25 30

40

Asn Gln Lys Ala Lys Glu Ala Leu Asp Val Ala Lys Lys Leu Gln Pro  
35 40 45

Ile Gln Thr Ser Ala Lys Asn Leu Ile Leu Phe Leu Gly Asp Gly Met  
 50 55 60

5 Gly Val Pro Thr Val Thr Ala Thr Arg Ile Leu Lys Gly Gln Leu Gly  
 65 70 75 80

Gly His Leu Gly Pro Glu Thr Pro Leu Ala Met Asp His Phe Pro Phe  
 85 90 95

10 Thr Ala Leu Ser Lys Thr Tyr Asn Val Asp Arg Gln Val Pro Asp Ser  
 100 105 110

Ala Gly Thr Ala Thr Ala Tyr Leu Cys Gly Val Lys Ala Asn Tyr Lys  
 115 120 125

15 Thr Ile Gly Val Ser Ala Ala Ala Arg Phe Asn Gln Cys Asn Ser Thr  
 130 135 140

20 Phe Gly Asn Glu Val Phe Ser Val Met His Arg Ala Lys Lys Ala Gly  
 145 150 155 160

Lys Ser Val Gly Val Val Thr Thr Thr Arg Val Gln His Ala Ser Pro  
 165 170 175

25 Ala Gly Thr Tyr Ala His Thr Val Asn Arg Asp Trp Tyr Ser Asp Ala  
 180 185 190

Asp Met Pro Ser Ser Ala Leu Gln Glu Gly Cys Lys Asp Ile Ala Thr  
 195 200 205

30 Gln Leu Ile Ser Asn Met Asp Ile Asp Val Ile Leu Gly Gly Gly Arg  
 210 215 220

Lys Phe Met Phe Pro Lys Gly Thr Pro Asp Pro Glu Tyr Pro Gly Asp  
 225 230 235 240

35 Ser Asp Gln Ser Gly Val Arg Leu Asp Ser Arg Asn Leu Val Glu Glu  
 245 250 255

Trp Leu Ala Lys Tyr Gln Gly Thr Arg Tyr Val Trp Asn Arg Glu Gln  
 260 265 270

40

34

	Leu	Met	Gln	Ala	Ser	Gln	Asp	Pro	Ala	Val	Thr	Arg	Leu	Met	Gly	Leu
			275					280					285			
5	Phe	Glu	Pro	Thr	Glu	Met	Lys	Tyr	Asp	Val	Asn	Arg	Asn	Ala	Ser	Ala
	290						295					300				
	Asp	Pro	Ser	Leu	Ala	Glu	Met	Thr	Glu	Val	Ala	Val	Arg	Leu	Leu	Ser
	305					310					315					320
10	Arg	Asn	Pro	Gln	Gly	Phe	Tyr	Leu	Phe	Val	Glu	Gly	Gly	Arg	Ile	Asp
					325					330					335	
	Gln	Gly	His	His	Ala	Gly	Thr	Ala	Tyr	Leu	Ala	Leu	Thr	Glu	Ala	Val
				340					345					350		
15	Met	Phe	Asp	Ser	Ala	Ile	Glu	Lys	Ala	Ser	Gln	Leu	Thr	Asn	Glu	Lys
			355					360					365			
	Asp	Thr	Leu	Thr	Leu	Ile	Thr	Ala	Asp	His	Ser	His	Val	Phe	Ala	Phe
	370						375					380				
20	Gly	Gly	Tyr	Thr	Leu	Arg	Gly	Thr	Ser	Ile	Phe	Gly	Leu	Ala	Pro	Leu
	385					390					395					400
	Asn	Ala	Gln	Asp	Gly	Lys	Ser	Tyr	Thr	Ser	Ile	Leu	Tyr	Gly	Asn	Gly
					405					410					415	
25	Pro	Gly	Tyr	Val	Leu	Asn	Ser	Gly	Asn	Arg	Pro	Asn	Val	Thr	Asp	Ala
				420					425					430		
30	Glu	Ser	Gly	Asp	Val	Asn	Tyr	Lys	Gln	Gln	Ala	Ala	Val	Pro	Leu	Ser
			435					440					445			
	Ser	Glu	Thr	His	Gly	Gly	Glu	Asp	Val	Ala	Ile	Phe	Ala	Arg	Gly	Pro
	450						455					460				
35	Gln	Ala	His	Leu	Val	His	Gly	Val	Gln	Glu	Gln	Asn	Tyr	Ile	Ala	His
	465					470					475					480
	Val	Met	Ala	Phe	Ala	Gly	Cys	Leu	Glu	Pro	Tyr	Thr	Asp	Cys	Gly	Leu
					485					490					495	
40	Ala	Pro	Pro	Ala	Asp	Glu	Asn	Arg	Pro	Thr	Thr	Pro	Val	Gln	Asn	Ser
				500					505					510		

35

Ala Ile Thr Met Asn Asn Val Leu Leu Ser Leu Gln Leu Leu Val Ser  
 515 520 525

5 Met Leu Leu Leu Val Gly Thr Ala Leu Val Val Ser  
 530 535 540

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:  
 10 (A) LENGTH: 559 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

15 Met Gln Gly Pro Trp Val Leu Leu Leu Leu Gly Leu Arg Leu Gln Leu  
 1 5 10 15

Ser Leu Ser Val Ile Pro Val Glu Glu Glu Asn Pro Ala Phe Trp Asn  
 20 25 30

20 Lys Lys Ala Ala Glu Ala Leu Asp Ala Ala Lys Lys Leu Gln Pro Ile  
 35 40 45

Gln Thr Ser Ala Lys Asn Leu Ile Ile Phe Leu Gly Asp Gly Met Gly  
 50 55 60

25 Val Pro Thr Val Thr Ala Thr Arg Ile Leu Lys Gly Gln Leu Glu Gly  
 65 70 75 80

His Leu Gly Pro Glu Thr Pro Leu Ala Met Asp Arg Phe Pro Tyr Met  
 85 90 95

30 Ala Leu Ser Lys Thr Tyr Ser Val Asp Arg Gln Val Pro Asp Ser Ala  
 100 105 110

Ser Thr Ala Thr Ala Tyr Leu Cys Gly Val Lys Thr Asn Tyr Lys Thr  
 115 120 125

35 Ile Gly Leu Ser Ala Ala Ala Arg Phe Asp Gln Cys Asn Thr Thr Phe  
 130 135 140

40 Gly Asn Glu Val Phe Ser Val Met Tyr Arg Ala Lys Lys Ala Gly Lys  
 145 150 155 160

36

	Ser	Val	Gly	Val	Val	Thr	Thr	Thr	Arg	Val	Gln	His	Ala	Ser	Pro	Ser
				165						170					175	
5	Gly	Thr	Tyr	Val	His	Thr	Val	Asn	Arg	Asn	Trp	Tyr	Gly	Asp	Ala	Asp
				180					185					190		
	Met	Pro	Ala	Ser	Ala	Leu	Arg	Glu	Gly	Cys	Lys	Asp	Ile	Ala	Thr	Gln
			195					200					205			
10	Leu	Ile	Ser	Asn	Met	Asp	Ile	Asn	Val	Ile	Leu	Gly	Gly	Gly	Arg	Lys
		210					215					220				
	Tyr	Met	Phe	Pro	Ala	Gly	Thr	Pro	Asp	Pro	Glu	Tyr	Pro	Asn	Asp	Ala
	225					230					235					240
15	Asn	Glu	Thr	Gly	Thr	Arg	Leu	Asp	Gly	Arg	Asn	Leu	Val	Gln	Glu	Trp
					245					250					255	
	Leu	Ser	Lys	His	Gln	Gly	Ser	Gln	Tyr	Val	Trp	Asn	Arg	Glu	Gln	Leu
				260					265					270		
20	Ile	Gln	Lys	Ala	Gln	Asp	Pro	Ser	Val	Thr	Tyr	Leu	Met	Gly	Leu	Phe
			275					280					285			
	Glu	Pro	Val	Asp	Thr	Lys	Phe	Asp	Ile	Gln	Arg	Asp	Pro	Leu	Met	Asp
25		290					295					300				
	Pro	Ser	Leu	Lys	Asp	Met	Thr	Glu	Thr	Ala	Val	Lys	Val	Leu	Ser	Arg
	305					310					315					320
30	Asn	Pro	Lys	Gly	Phe	Tyr	Leu	Phe	Val	Glu	Gly	Gly	Arg	Ile	Asp	Arg
					325					330					335	
	Gly	His	His	Leu	Gly	Thr	Ala	Tyr	Leu	Ala	Leu	Thr	Glu	Ala	Val	Met
				340					345					350		
35	Phe	Asp	Leu	Ala	Ile	Glu	Arg	Ala	Ser	Gln	Leu	Thr	Ser	Glu	Arg	Asp
			355					360					365			
	Thr	Leu	Thr	Ile	Val	Thr	Ala	Asp	His	Ser	His	Val	Phe	Ser	Phe	Gly
		370					375					380				

40



37

Gly Tyr Thr Leu Arg Gly Thr Ser Ile Phe Gly Leu Ala Pro Leu Asn  
 385 390 395 400  
 Ala Leu Asp Gly Lys Pro Tyr Thr Ser Ile Leu Tyr Gly Asn Gly Pro  
 405 410 415  
 5  
 Gly Tyr Val Gly Gly Thr Gly Glu Arg Pro Asn Val Thr Ala Ala Glu  
 420 425 430  
 10 Ser Ser Gly Ser Ser Tyr Arg Arg Gln Ala Ala Val Pro Val Lys Ser  
 435 440 445  
 Glu Thr His Gly Gly Glu Asp Val Ala Ile Phe Ala Arg Gly Pro Gln  
 450 455 460  
 15 Ala His Leu Val His Gly Val Gln Glu Gln Asn Tyr Ile Ala His Val  
 465 470 475 480  
 Met Ala Ser Ala Gly Cys Leu Glu Pro Tyr Thr Asp Cys Gly Leu Ala  
 485 490 495  
 20 Pro Pro Ala Asp Glu Ser Gln Thr Thr Thr Thr Thr Arg Gln Thr Thr  
 500 505 510  
 Ile Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Pro Val His  
 515 520 525  
 25 Asn Ser Ala Arg Ser Leu Gly Pro Ala Thr Ala Pro Leu Ala Leu Ala  
 530 535 540  
 30 Leu Leu Ala Gly Met Leu Met Leu Leu Leu Gly Ala Pro Ala Glu  
 545 550 555

## (2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 528 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear  
 35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Gln Gly Pro Trp Val Leu Leu Leu Leu Gly Leu Arg Leu Gln Leu  
 1 5 10 15  
 40 Ser Leu Gly Val Ile Pro Ala Glu Glu Glu Asn Pro Ala Phe Trp Asn  
 20 25 30

Arg Gln Ala Ala Glu Ala Leu Asp Ala Ala Lys Lys Leu Gln Pro Ile  
 35 40 45

5 Gln Lys Val Ala Lys Asn Leu Ile Leu Phe Leu Gly Asp Gly Leu Gly  
 50 55 60

Val Pro Thr Val Thr Ala Thr Arg Ile Leu Lys Gly Gln Lys Asn Gly  
 65 70 75 80

10 Lys Leu Gly Pro Glu Thr Pro Leu Ala Met Asp Arg Phe Pro Tyr Leu  
 85 90 95

Ala Leu Ser Lys Thr Tyr Asn Val Asp Arg Gln Val Pro Asp Ser Ala  
 100 105 110

15 Ala Thr Ala Thr Ala Tyr Leu Cys Gly Val Lys Ala Asn Phe Gln Thr  
 115 120 125

20 Ile Gly Leu Ser Ala Ala Ala Arg Phe Asn Gln Cys Asn Thr Thr Arg  
 130 135 140

Gly Asn Glu Val Ile Ser Val Met Asn Arg Ala Lys Gln Ala Gly Lys  
 145 150 155 160

25 Ser Val Gly Val Val Thr Thr Thr Arg Val Gln His Ala Ser Pro Ala  
 165 170 175

Gly Thr Tyr Ala His Thr Val Asn Arg Asn Trp Tyr Ser Asp Ala Asp  
 180 185 190

30 Met Pro Ala Ser Ala Arg Gln Glu Gly Cys Gln Asp Ile Ala Thr Gln  
 195 200 205

Leu Ile Ser Asn Met Asp Ile Asp Val Ile Leu Gly Gly Gly Arg Lys  
 210 215 220

35 Tyr Met Phe Pro Met Gly Thr Pro Asp Pro Glu Tyr Pro Ala Asp Ala  
 225 230 235 240

Ser Gln Asn Gly Ile Arg Leu Asp Gly Lys Asn Leu Val Gln Glu Trp  
 245 250 255

40

39

Leu Ala Lys His Gln Gly Ala Trp Tyr Val Trp Asn Arg Thr Glu Leu  
 260 265 270

5 Met Glu Ala Ser Leu Asp Gln Ser Val Thr His Leu Met Gly Leu Phe  
 275 280 285

Glu Pro Gly Asp Thr Lys Tyr Glu Ile His Arg Asp Pro Thr Leu Asp  
 290 295 300

10 Pro Ser Leu Met Glu Met Thr Glu Ala Ala Leu Arg Leu Leu Ser Arg  
 305 310 315 320

Asn Pro Arg Gly Phe Tyr Leu Phe Val Glu Gly Gly Arg Ile Asp His  
 325 330 335

15 Gly His His Glu Gly Val Ala Tyr Gln Ala Leu Thr Glu Ala Val Met  
 340 345 350

Phe Asp Asp Ala Ile Glu Arg Ala Gly Gln Leu Thr Ser Glu Glu Asp  
 355 360 365

20 Thr Leu Thr Leu Val Thr Ala Asp His Ser His Val Phe Ser Phe Gly  
 370 375 380

25 Gly Tyr Thr Leu Arg Gly Ser Ser Ile Phe Gly Leu Ala Pro Ser Lys  
 385 390 395 400

Ala Gln Asp Ser Lys Ala Tyr Thr Ser Thr Leu Tyr Gly Asn Gly Pro  
 405 410 415

30 Gly Tyr Val Phe Asn Ser Gly Val Arg Pro Asp Val Asn Glu Ser Glu  
 420 425 430

Ser Gly Ser Pro Asp Tyr Gln Gln Gln Ala Ala Val Pro Leu Ser Ser  
 435 440 445

35 Glu Thr His Gly Gly Glu Asp Val Ala Val Phe Ala Arg Gly Pro Gln  
 450 455 460

Ala His Leu Val His Gly Val Gln Glu Gln Ser Phe Val Ala His Val  
 465 470 475 480

40 Met Ala Phe Ala Ala Cys Leu Glu Pro Tyr Thr Ala Cys Asp Leu Ala  
 485 490 495

40

Pro Pro Ala Cys Thr Thr Asp Ala Ala His Pro Val Ala Ala Ser Leu  
500 505 510

5 Pro Leu Leu Ala Gly Thr Leu Leu Leu Leu Gly Ala Ser Ala Ala Pro  
515 520 525

I claim:

1. An isolated nucleic acid comprising a nucleotide sequence encoding a substantially purified calf intestinal alkaline phosphatase or an active fragment thereof.
- 5 2. The isolated nucleic acid of claim 1 having a nucleotide sequence substantially the same as the nucleotide sequence of Figure 1.
3. The isolated nucleic acid of claim 1, wherein said nucleic acid is cDNA.
- 10 4. The isolated nucleic acid of claim 1, wherein said nucleic acid is RNA.
5. The isolated nucleic acid of claim 1, further comprising a second nucleotide sequence encoding a polypeptide having specific reactivity with a ligand.
- 15 6. A vector comprising the nucleic acid of claim 1.
7. The vector of claim 6, wherein said vector is a plasmid.
8. A recombinant host cell comprising the  
20 vector of claim 6.
9. A recombinant polypeptide produced by the recombinant host cell of claim 8.

10. A method of obtaining recombinant calf intestinal alkaline phosphatase or an active fragment thereof, comprising culturing said recombinant host cell of claim 8 and isolating said calf intestinal alkaline  
5 phosphatase or active fragment thereof from said culture.

11. A cell culture comprising the recombinant host cell of claim 8 cultured in a suitable medium.

12. A nucleic acid probe comprising a nucleotide sequence complementary to a portion of a nucleotide  
10 sequence specific to calf intestinal alkaline phosphatase.

13. A multifunctional polypeptide comprising an amino acid sequence of calf intestinal alkaline phosphatase or an active fragment thereof and a second amino acid sequence of a reagent having specific reactivity with a  
15 desired ligand.

14. The multifunctional polypeptide of claim 13, wherein said reagent encoded by the second amino acid sequence is an antibody.

15. A method for determining the presence of a  
20 ligand in a sample, comprising:

(a) contacting said sample with a substantially purified calf intestinal alkaline phosphatase and a reagent that specifically binds to said ligand, said reagent attached to said recombinant calf intestinal alkaline  
25 phosphatase;

(b) contacting said sample with a detectable substrate catalyzed by the recombinant polypeptide; and

(c) detecting the binding of said sample to the reagent, wherein binding indicates the presence of said  
30 ligand in the sample.



16. The method of claim 15, further comprising the step of (d) determining an amount of binding of said sample to the reagent, wherein the amount of binding relates to the concentration of said ligand in the sample.

5 17. The method of claim 15, wherein said reagent is an anti-ligand antibody.

18. The method of claim 15, wherein said reagent and recombinant calf IAP or active fragment thereof are attached as a multifunctional polypeptide.

10 19. The method of claim 15, wherein said reagent is an oligonucleotide.

20. The method of claim 19, wherein said ligand is a cDNA or genomic DNA fragment.

## AMENDED CLAIMS

[received by the International Bureau on 9 July 1993 (09.07.93);  
original claims 1-20 replaced by amended claims 1-25 (4 pages)]

1. An isolated nucleic acid comprising:
  - (a) the nucleotide sequence shown in Figure 1 encoding calf intestinal alkaline phosphatase;
  - (b) substantially the same nucleotide
  - 5 sequence as the sequence shown in Figure 1, encoding calf intestinal alkaline phosphatase; or
  - (c) a nucleotide sequence encoding an active fragment of a calf intestinal alkaline phosphatase encoded by a portion of a nucleotide sequence of (a) or
  - 10 (b).
2. The isolated nucleic acid of claim 1, wherein the nucleotide sequence is the coding sequence shown in Figure 1.
3. An isolated nucleic acid sequence,
- 15 comprising a nucleotide sequence encoding the amino acid sequence of calf intestinal alkaline phosphatase of Figure 1.
4. The nucleic acid of claim 1 wherein the nucleic acid is cDNA.
- 20 5. An isolated RNA molecule encoding the amino acid sequence of calf intestinal alkaline phosphatase of Figure 1, or an active fragment of the calf intestinal alkaline phosphatase of Figure 1.
6. The isolated nucleic acid of claim 1,
- 25 further comprising a second nucleotide sequence encoding a polypeptide having specific reactivity with a ligand.
7. A vector comprising the isolated nucleic acid of claim 1.

8. The vector of claim 7, wherein the vector is a plasmid.
9. A host cell comprising the vector of claim 7.
- 5 10. A recombinant polypeptide produced by the host cell of claim 9.
- 10 11. A method of obtaining recombinant calf intestinal alkaline phosphatase or an active fragment thereof, comprising culturing the host cell of claim 9 and isolating the calf intestinal alkaline phosphatase or active fragment thereof from the culture.
12. A cell culture comprising the host cell of claim 9 and a suitable medium.
- 15 13. A nucleic acid probe comprising a nucleotide sequence complementary to a portion of the nucleotide sequence of the coding region of the sequence shown in Figure 1.
- 20 14. A composition comprising recombinant calf intestinal alkaline phosphatase or an active fragment thereof attached to a reagent specifically reactive to a ligand to be detected.
- 25 15. The composition of claim 14, wherein the alkaline phosphatase or an active fragment thereof attached to a reagent comprises a multifunctional polypeptide.
16. The composition of claim 14, wherein the alkaline phosphatase or an active fragment thereof is chemically coupled to the reagent.

17. The composition of any of claims 14-16, wherein the reagent comprises an antibody or a reactive fragment thereof.

18. The composition of claim 17, wherein the  
5 reagent has specific reactivity with a cancer marker, allergen or drug.

19. The composition of claim 17, wherein the reagent has specific reactivity with a nucleic acid.

20. A method for determining the presence of a  
10 ligand in a sample, comprising:

(a) contacting the sample with recombinant calf intestinal alkaline phosphatase or an active fragment thereof, wherein the recombinant calf intestinal alkaline phosphatase or an active fragment is attached to  
15 a reagent specifically reactive with said ligand;

(b) contacting the sample with a detectable agent catalyzed by calf intestinal alkaline phosphatase; and

(c) detecting the binding of the sample to the  
20 reagent, wherein binding indicates the presence of said ligand in the sample.

21. The method of claim 20, further comprising the step of:

(d) relating the amount of binding to the  
25 concentration of the ligand.

22. The method of claim 20, wherein the reagent is an anti-ligand antibody.

23. The method of claim 20, wherein the reagent and recombinant calf intestinal alkaline  
30 phosphatase or active fragment thereof are attached as a multifunctional polypeptide.

24. The method of claim 20, wherein the reagent specifically reacts with an oligonucleotide.

25. The method of claim 24, wherein the reagent specifically reacts with a cDNA or genomic DNA  
5 fragment.

STATEMENT UNDER ARTICLE 19

Amended claims 1-3 and 5 find support on page 5.  
Amended claim 13 finds support on page 8, lines 25-33.  
Amended claims 14-18 find support on page 8, lines 8-9 and 20-23, page 9, lines 22-25, and page 10, lines 15-22. Amended claims 20 and 21 find support on page 9, lines 11-34.

Other amendments, such as replacing "said" with --the--, are clerical in nature.



FIGURE 1A

1/11

AAGCTTTCAC CTTCTCTGAA AACAGAGAGA CAGTCCTCAG CCCCAGTCCT CACCCTTCCT	60
ACCTCCCTGC CTGATGCCCA GGCAATCATC TGGTGGCGTG TCACCTCCCT CTGTCCCATG	120
AGTTCCACTA GATGTGGCCC TCAAGAAAAA GGGCTTCCCT GTTGGCTCAG CTGGTAAAGA	180
ATCCTCCAGC AATGTAGGAG ACCTGGGTTC GATCCCTGGG TTGGGAGGAT ACCCTGGAGA	240
AGGGAATGGC TACCCACTCC AGTATTCTTG CCTGGATAAT CCCATGGACA GAGGAGTCTG	300
GCAGGCTGCA GACCATAAGG TAGAAAGAGT CAGACATGAC TGAGCAACTA AGCACAATAT	360
TCCACTGGAT ATATCATACT TTGTTTCATCC ATTTGTCTGC TGTGGATGGT TGAGTGGCTT	420
GTGCCTCTTG GCTACTGTGA GTAATGCTAC TAAATGTGA GTGTGCAAAT ACCTCTTATA	480
GATCTTGATT TCAATTATTG GGGATACACA CCCAGAAGGC GGATTGTTGG ATGTGAGAAT	540
GCCTTTTGA ACCCCAACCT GGGGTACTG AAACCCTAGC TCCTTATCAG AAGCTGTTCC	600
TGTGAGTGTG TGTGGCCTGT GGAGAGAAGA GACTCACCTC TGCCTTCCAT TTACCTCTCC	660
AATGGAGCAG AGGTTGCAAA CTTCAAGTTAA TGGGCACTGG GCCCAGCCTT GTCGACCCGT	720
TACAGGCACC TTACACACAC ACACACACAC ACACACACAC ACAACAGCA CTGCAGACCC	780
AGCTCTTCAG TAACTGAAGA CACAGACAAG GCCCCCGCTC TGCTGTCACC TCCAGTCCCA	840
TCCTTCTCCA CAGCAGAAGC TGGGCCCAGG CTCCCATGTG CCCCCTAG CCCAGTGCCC	900
ACACCTCCTG CCCAGGTCAA GTCTGGTGAG GAGCTGAGCA GGGGGCAGGG CAGACAGGCC	960
TCCCCGTGGA TCTCTGTCTC AGGGCGCCAG GGAACCTAACC CAGGCCCTG GCCAGGCTGT	1020
GTCCCTAAGC ACTGGGAACC AAACCAGGCC AAGGCTGAGT CTCAGAAAAC ACTGAACACG	1080
TGAAGGAAGG AGAGATGGTT CTCCACAGG ACTTGGTGAG CAGAGGGCTG GGAGGAGCCT	1140
CAGTCAGGAC CTTGAAAACG TTCCTCAGGC CTAGACATCT GCACCCTAAT CCCCACCCCA	1200
CCCTGAGGAG ACAGCTGGGA CCATCCTGGG AGGGAGGGAC CTGAATCCTC AGGACCCCTA	1260
CTGCTAAGCC ACACCCACCA CATGCCCCCTG GCAACAGGGC TCAAAGTCAT AGGGCAGGTG	1320
AGGGGCAGGG TGTGGCCACC CGGGGAACCT GGGATGGACA AGGAGACTTT AATAGCAGGG	1380
ACAAAGTCTA TCTAGATTTA AGCCCAGCAG GCCAAGCTGC AGCCGGTCCC TGGTGTCCCA	1440
GCCTTGCCCT GAGACCCGGC CTCCCCAGGT CCCATCCTGA CCCTCTGCCA TCACACAGCC	1500
ATG CAG GGG GCC TGC GTG CTG CTG CTG CTG GGC CTG CAT CTA CAG CTC	1548
Met Gln Gly Ala Cys Val Leu Leu Leu Leu Gly Leu His Leu Gln Leu	
1 5 10 15	
TCC CTA GGC CTC GTC CCA G GTAATCAGGC GGCTCCCAGC AGCCCTACT	1597
Ser Leu Gly Leu Val Pro	
20	
CACAGGGGCG GCTCTAGGCT GACCTGACCA AACTCTCCC CTTGGGCAG TT GAG	1651
Val Glu	
GAG GAA GAC CCC GCC TTC TGG AAC CGC CAG GCA GCC CAG GCC CTC GAT	1699
Glu Glu Asp Pro Ala Phe Trp Asn Arg Gln Ala Ala Gln Ala Leu Asp	
25 30 35 40	

SUBSTITUTE SHEET

Figure 1B

2/11

GTG GCT AAG AAG CTG CAG CCC ATC CAG ACA GCC GCC AAG AAT GTC ATC Val Ala Lys Lys Leu Gln Pro Ile Gln Thr Ala Ala Lys Asn Val Ile 45 50 55	1747
CTC TTC TTG GGG GAT G GTGAGTACAT GAGGCCAGCC CACCCCCTGT Leu Phe Leu Gly Asp 60	1793
CCCCTGACAG GCCTGGAACC CTGTGATGCC GGCTGACCCA GGTGCGCCCC AGAAACTCGG	1853
ACCTGAGACA CTGTGTACCT TCAG GG ATG GGG GTG CCT ACG GTG ACA GCC Gly Met Gly Val Pro Thr Val Thr Ala 65 70	1903
ACT CGG ATC CTA AAG GGG CAG ATG AAT GGC AAA CTG GGA CCT GAG ACA Thr Arg Ile Leu Lys Gly Gln Met Asn Gly Lys Leu Gly Pro Glu Thr 75 80 85	1951
CCC CTG GCC ATG GAC CAG TTC CCA TAC GTG GCT CTG TCC AAG Pro Leu Ala Met Asp Gln Phe Pro Tyr Val Ala Leu Ser Lys 90 95 100	1993
GTAAGGCCAA GTGGCCTCAG GGTGGTCTAC ACCAGAGGGG TGGGTGTGGG CCTAGGGAGC	2053
AGGGTAGGAG GGAAACCCAG GAGGGCTAGG GGCTGAGATA GGGGCTGGGG GCTGTGAGGA	2113
TGGGCCCAGG GCTGGGTCAG GAGCTGGGTG TCTACCCAGC AGAGCGTAAG GCATCTCTGT	2173
CCCAG ACA TAC AAC GTG GAC AGA CAG GTG CCA GAC AGC GCA GGC ACT Thr Tyr Asn Val Asp Arg Gln Val Pro Asp Ser Ala Gly Thr 105 110	2220
GCC ACT GCC TAC CTG TGT GGG GTC AAG GGC AAC TAC AGA ACC ATT GGT Ala Thr Ala Tyr Leu Cys Gly Val Lys Gly Asn Tyr Arg Thr Ile Gly 115 120 125 130	2268
GTA AGT GCA GCC GCC CGC TAC AAC CAG TGC AAA ACG ACA CGT GGG AAT Val Ser Ala Ala Ala Arg Tyr Asn Gln Cys Lys Thr Thr Arg Gly Asn 135 140 145	2316
GAG GTC ACG TCT GTG ATG AAC CGG GCC AAG AAA GCA G GTGGGCTTGG Glu Val Thr Ser Val Met Asn Arg Ala Lys Lys Ala 150 155	2363
GCGTCAGCTT CCTGGGCAGG GACGGGCTCA GAGACCTCAG TGGCCCACCG TGACCTCTGC	2423
CACCCTCAG GG AAG TCC GTG GGA GTG GTG ACC ACC ACC AGG GTG CAG Gly Lys Ser Val Gly Val Val Thr Thr Thr Arg Val Gln 160 165 170	2470
CAT GCC TCC CCA GCC GGG GCC TAC GCG CAC ACG GTG AAC CGA AAC TGG His Ala Ser Pro Ala Gly Ala Tyr Ala His Thr Val Asn Arg Asn Trp 175 180 185	2518
TAC TCA GAC GCC GAC CTG CCT GCT GAT GCA CAG ATG AAT GGC TGC CAG Tyr Ser Asp Ala Asp Leu Pro Ala Asp Ala Gln Met Asn Gly Cys Gln 190 195 200	2566
GAC ATC GCC GCA CAG CTG GTC AAC AAC ATG GAT ATT GAC GTGCGACATG Asp Ile Ala Ala Gln Leu Val Asn Asn Met Asp Ile Asp 205 210 215	2615
TTGGGCACAG GGCGGGGCTG GGCACAGGTG GTGGGGCACA CTCGCAACAC AGTCGTAGGT	2675
AACCTCCAGC CTGCGGTGTT TCAGGGTTTT CATGGGTTTG TGTGTGTGTG TATGTGTGGT	2735
GGGGTGGCAC CATGTAGGAG GTGGGGACAG GCCTTTCCCA CAGACCTGGT GGGGGAGGTA	2795

SUBSTITUTE SHEET

Figure 1c 3/11

GGGGCTGTGT GAGAGGAGTA AAGGGCCAGC CAGGCCCTA ACCCACCTGC CTAACCTCT	2855
GGCTCCAG GTG ATC CTG GGT GGA GGC CGA AAA TAC ATG TTT CCT GTG GGG Val Ile Leu Gly Gly Gly Arg Lys Tyr Met Phe Pro Val Gly 220 225 230	2905
ACC CCA GAC CCT GAA TAC CCA GAT GAT GCC AGT GTG AAT GGA GTC CGG Thr Pro Asp Pro Glu Tyr Pro Asp Asp Ala Ser Val Asn Gly Val Arg 235 240 245	2953
AAG CGA AAG CAG AAC CTG GTG CAG GCA TGG CAG GCC AAG CAC CAG Lys Arg Lys Gln Asn Leu Val Gln Ala Trp Gln Ala Lys His Gln 250 255 260	2998
GTAATGGGGG CTCACGGATG TGGGGGTACA GTGGGGCTGG GCCTGGGGTG TCGGCTATGG	3058
CTGAGGCCTG GTTCTGCCCT CCCAG GGA GCC CAG TAT GTG TGG AAC CGC ACT Gly Ala Gln Tyr Val Trp Asn Arg Thr 265 270	3110
GCG CTC CTT CAG GCG GCC GAT GAC TCC AGT GTA ACA CAC CTC ATG G Ala Leu Leu Gln Ala Ala Asp Asp Ser Ser Val Thr His Leu Met 275 280 285	3156
GTAACGACTC CACCCACCCT CACTGTCCTC CCCAGGAATG GGTGCCATGG GCCACCCCTG	3216
TCCTCAGCTT GAGGGTCACC ACTGCTCCCC TTTCCCACAG GC CTC TTT GAG CCG Gly Leu Phe Glu Pro 290	3270
GCA GAC ATG AAG TAT AAT GTT CAG CAA GAC CAC ACC AAG GAC CCG ACC Ala Asp Met Lys Tyr Asn Val Gln Gln Asp His Thr Lys Asp Pro Thr 295 300 305	3318
CTG CAG GAA ATG ACA GAG GTG GCC CTG CGA GTC GTA AGC AGG AAC CCC Leu Gln Glu Met Thr Glu Val Ala Leu Arg Val Val Ser Arg Asn Pro 310 315 320	3366
AGG GGC TTC TAC CTC TTT GTG GAG G GTGAGTGGCA GCCCCTTGGT Arg Gly Phe Tyr Leu Phe Val Glu 325 330	3411
GAACAGAGGT GTGATGAGGG CCATCAGGGT GGGTTTGGTA TCTTATATGT GACTTATCTG	3471
CAG GA GGC CGC ATT GAC CAC GGT CAC CAT GAT GAC AAA GCT TAT ATG Gly Gly Arg Ile Asp His Gly His His Asp Asp Lys Ala Tyr Met 335 340 345	3518
GCA CTG ACC GAG GCG GGT ATG TTT GAC AAT GCC ATC GCC AAG GCT AAT Ala Leu Thr Glu Ala Gly Met Phe Asp Asn Ala Ile Ala Lys Ala Asn 350 355 360	3566
GAG CTC ACT AGC GAA CTG GAC ACG CTG ATC CTT GTC ACT GCA GAC CAC Glu Leu Thr Ser Glu Leu Asp Thr Leu Ile Leu Val Thr Ala Asp His 365 370 375	3614
TCT CAT GTC TTC TCT TTT GGT GGC TAT ACA CTG CGT GGG ACC TCC ATT Ser His Val Phe Ser Phe Gly Gly Tyr Thr Leu Arg Gly Thr Ser Ile 380 385 390	3662
TTT G GTAAGCCCAG GGAGAGTGGC AGGTCGTTGC CCCTAAGTTA CGAGGCACAA Phe	3716
CTCGTCTGAG CCAGTTCCTC TATCTGTCTA GTGGGGTAGT ACAGCACACT GCCTGCTACG	3776
CTCTGGTGAG GATTGTCACT GACAGACAGA CTGGCCATGG CTCTGCACAC AGGGGAGCAC	3836

SUBSTITUTE SHEET

Figure 1D 4/11

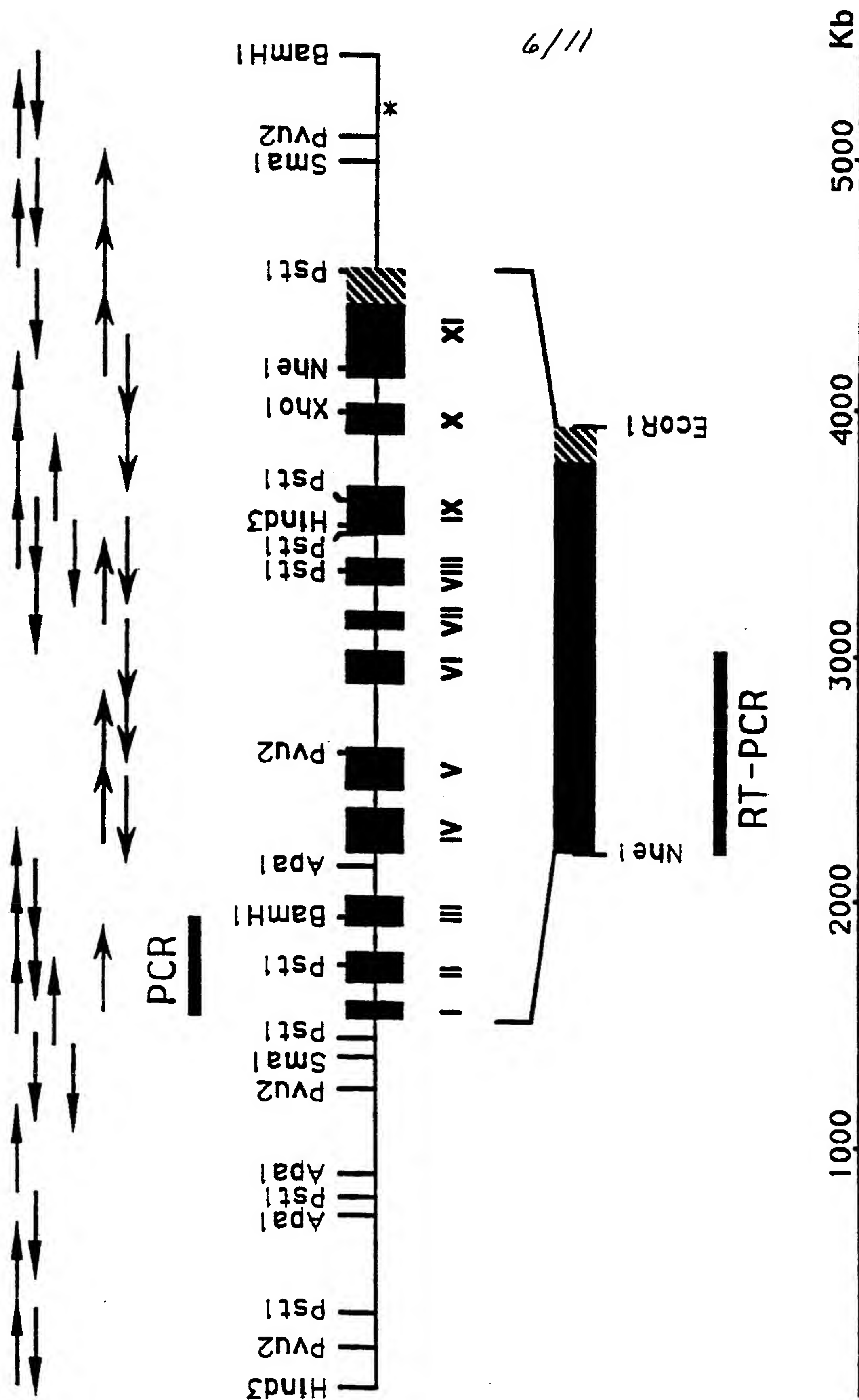
AAGCTAGGTC. AGTGTGATCA CGGGGTCCCC TCTTCCCTGA AG GT CTG GCC CCC Gly Leu Ala Pro 395	3889
AGC AAG GCC TTA GAC AGC AAG TCC TAC ACC TCC ATC CTC TAT GGC AAT Ser Lys Ala Leu Asp Ser Lys Ser Tyr Thr Ser Ile Leu Tyr Gly Asn 400 405 410	3937
GGC CCA GGC TAT GCG CTT GGC GGG GGC TCG AGG CCC GAT GTT AAT GAC Gly Pro Gly Tyr Ala Leu Gly Gly Gly Ser Arg Pro Asp Val Asn Asp 415 420 425 430	3985
AGC ACA. AGC G GTAAGTGTAG TAGGTGGGGC GCTGGGAGGT GGGGACCCTG Ser Thr Ser	4035
GCCAGAAATT GTGGGGAGGG GAAGGCTGCC TCCCTTGTC CATTAACTTC CCTTCTTCTG	4095
GCCAG AG GAC CCC TCG TAC CAG CAG CAG GCG GCC GTG CCC CAG GCT Glu Asp Pro Ser Tyr Gln Gln Gln Ala Ala Val Pro Gln Ala 435 440 445	4141
AGC GAG ACC CAC GGG GGC GAG GAC GTG GCG GTG TTC GCG CGC GGC CCG Ser Glu Thr His Gly Gly Glu Asp Val Ala Val Phe Ala Arg Gly Pro 450 455 460	4189
CAG GCG CAC CTG GTG CAC GGC GTC GAG GAG GAG ACC TTC GTG GCG CAC Gln Ala His Leu Val His Gly Val Glu Glu Glu Thr Phe Val Ala His 465 470 475	4237
ATC ATG GCC TTT GCG GGC TGC GTG GAG CCC TAC ACC GAC TGC AAT CTG Ile Met Ala Phe Ala Gly Cys Val Glu Pro Tyr Thr Asp Cys Asn Leu 480 485 490 495	4285
CCA GCC CCC ACC ACC GCC ACC AGC ATC CCC GAC GCC GCG CAC CTG GCG Pro Ala Pro Thr Thr Ala Thr Ser Ile Pro Asp Ala Ala His Leu Ala 500 505 510	4333
GCC AGC CCG CCT CCA CTG GCG CTG CTG GCT GGG GCG ATG CTG CTG CTG Ala Ser Pro Pro Pro Leu Ala Leu Leu Ala Gly Ala Met Leu Leu Leu 515 520 525	4381
CTG GCG CCC ACC TTG TAC TAACCCCCAC CAGTTCCAGG TCTCGGGATT Leu Ala Pro Thr Leu Tyr 530	4429
TCCCGCTCTC CTGCCCCAAA CCTCCCAGCT CAGGCCCTAC CGGAGCTACC ACCTCAGAGT	4489
CCCCACCCCG AAGTGCTATC CTAGCTGCCA CTCCTGCAGA CCCGACCCGG CCCCACCACC	4549
AGAGTTTCAC CTCCCAGCAG TGATTACAT TCCAGCATTG AAGGAGCCTC AGCTAACAGC	4609
CCTTCAAGGC CCAGCCTATA CCGGAGGCTG AGGCTCTGAT TTCCCTGTGA CACGCGTAGA	4669
CCTACTGCCC GACCCCAACT TCGGTGGCTT GGGATTTTGT GTTCTGCCAC CCTGAACCTC	4729
AGTAAGGGGG CTCGGACCAT CCAGACTGCC CCTACTGCCC ACAGCCCACC TGAGGACAAA	4789
GCTGGCACGG TCCCAGGGGT CCCAGGCCCG GCTGGAACCC ACACCTTGCC TTCAGCGACC	4849
TGGACTCTGG GTTCGGAGAG TGGCTTCGGG AGGCGTGGTT TCCGATGGGC GTGCTCTGGA	4909
ACGTGCTCGC CTGAACCAAC CTGTGTACAC TGGCCAGGAA TCACGGCCAC CAGAGCTCGG	4969
ACCTGACAGA GCCCTCAGCA GCCCCTCCTA GACCAACGTA CCCATTACAG AGAGGAGACA	5029
GAGACACAGA GGAGAGGAGA CTTGTCCCAG GTCCCTCAGC TGCTGTGAGG GCGGCCCTGG	5089

SUBSTITUTE SHEET

Figure 1E 5/11

TGCCCCCTTCC AGGCTGGGCA TCCAGTAGC AGCAGGGGAC CCGGGGGTGG GGACACAGGC	5149
CCCGCCCTCC CTGGGAGGCA GGAAGCAGCT CTCAAATAAA CTGTTCTAAG TATGATACAG	5209
GAGTGATACA TGTGTGAAGA GAAGCCCTTA GGTGGGGGCA CAGAGTGTCT GGGTGAGGGG	5269
GGTCAGGGTC ACATCAGGAG GTTAGGGAGG GGTGATGAA GGGCTGACGT TGAGCAAAGA	5329
CCAAAGGCAA CTCAGAAGGA CAGTGGTGCA GGA CTGGGTG TGGTCAGCAG GGGGACTGGT	5389
TGGGGGATCC	5399





## Figure 2



7/11

-19 M Q G A C V L L L L \* G L H L Q L S L G L V P V E E E D P A F  
-19 - - - P W - - - - - \* - - - R - - - - - V I - A - - - N - - -  
-20 - - - D W - - - - - L - - - R I H - - - F - V I - - - - - N - V -  
-19 - - - P W - - - - - \* - - - R - - - - - S V I - - - - - N - - -

G M G V P T V T A T R I L K G Q M N G K L G P E T P L A M D Q  
- L - - - - - - - - - - - K - - - - - - - - - - - R  
- - - - - - - - - - - - - L G - H - - - - - - - H  
- - - - - - - - - - - - - L E - H - - - - - - - R

G N Y R T I G V S A A A R Y N Q C K T T R G N E V T S V M N R  
A - F Q - - - L - - - - - F - - - - - N - - - - - I - - - - -  
A - - K - - - - - - - - F - - - - - N S - F - - - - - F - - - - - H -  
T - - K - - - L - - - - - F D - - - - - N - - - - - F - - - - - Y -

N W Y S D A D L P A D A Q M N G C Q D I A A Q L V N N M D I D  
- - - - - M - - - S - R Q E - - - - - T - - - I S - - - - -  
D - - - - - M - S S - L Q E - - K - - - - T - - - I S - - - - -  
- - - G - - - M - - S - L R E - - K - - - - T - - - I S - - - - - N

R K Q N L V Q A W Q A K H Q G A Q Y V W N R T A L L Q A A D D  
D G K - - - - E - L - - - - - W - - - - - E - M E - S L -  
D S R - - - - E E - L - - - Y - - - T R - - - - - E Q - M - - S Q -  
D G R - - - - E - L S - - - - - S - - - - - E Q - I - K - Q -

FIG.3A

8/11

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| W | N | R | Q | A | A | Q | A | L | D | V | A | K | K | L | Q | P | I | Q | T | A | A | K | N | V | I | L | F | L | G | D |
| - | - | - | - | - | - | E | - | - | - | A | - | - | - | - | - | - | - | - | - | K | V | - | - | L | - | - | - | - | - |   |
| - | - | Q | K | - | K | E | - | - | - | - | - | - | - | - | - | - | - | - | - | S | - | - | L | - | - | - | - | - |   |   |
| - | - | K | K | - | - | E | - | - | - | A | - | - | - | - | - | - | - | - | - | S | - | - | L | - | I | - | - | - |   |   |

42 b.IAP  
42 h.IAP  
42 r.IAP  
42 m.IAP

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| F | P | Y | V | A | L | S | K | T | Y | N | V | D | R | Q | V | P | D | S | A | G | T | A | T | A | Y | L | C | G | V | K |
| - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | A | - | - | - | - | - | - | - | - |   |   |
| - | - | F | T | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | G | - | - | - | - | - | - | - | - | - |   |   |
| - | - | M | - | - | - | - | - | - | - | S | - | - | - | - | - | - | - | - | S | - | - | - | - | - | - | - | - | - |   |   |

104 b.IAP  
104 h.IAP  
104 r.IAP  
104 m.IAP

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | K | K | A | G | K | S | V | G | V | T | T | T | R | V | Q | H | A | S | P | A | G | A | Y | A | H | T | V | N | R |
| - | - | Q | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | T | - | - | - | - | - | - | - |   |
| - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | T | - | - | - | - | - | - | - |   |   |
| - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | S | - | T | - | V | - | - | - |   |   |

166 b.IAP  
166 h.IAP  
166 r.IAP  
166 m.IAP

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| V | I | L | G | G | R | K | Y | M | F | P | V | G | T | P | D | P | E | Y | P | D | D | A | S | V | N | G | V | R | K |
| - | - | - | - | - | - | - | - | - | - | - | M | - | - | - | - | - | - | - | - | A | - | - | - | Q | - | - | I | - |   |
| - | - | - | - | - | - | - | F | - | - | - | K | - | - | - | - | - | - | - | G | - | S | D | Q | S | - | - | - |   |   |
| - | - | - | - | - | - | - | - | - | - | - | A | - | - | - | - | - | - | - | N | - | - | N | E | T | - | T | - |   |   |

228 b.IAP  
228 h.IAP  
228 r.IAP  
228 m.IAP

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S | S | V | T | H | L | M | G | L | F | E | P | A | D | M | K | Y | N | V | Q | Q | D | H | T | K | D | P | T | L | Q | E |
| Q | - | - | - | - | - | - | - | - | - | - | G | - | T | - | - | E | I | H | R | - | P | - | L | - | - | S | - | M | - |   |
| P | A | - | - | R | - | - | - | - | - | - | T | E | - | - | - | D | - | N | R | N | A | S | A | - | - | S | - | A | - |   |
| P | - | - | - | Y | - | - | - | - | - | - | V | - | T | - | F | D | I | - | R | - | P | L | M | - | - | S | - | K | D |   |

290 b.IAP  
290 h.IAP  
290 r.IAP  
290 m.IAP

FIG.3B

9/11

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| M | T | E | V | A | L | R | V | S | R | N | P | R | G | F | Y | L | F | V | E | G | G | R | I | D | H | G | H | D |
| - | - | - | A | - | - | - | L | L | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | E |   |
| - | - | - | - | V | - | L | L | - | - | - | - | Q | - | - | - | - | - | - | - | - | - | - | - | Q | - | - | A |   |
| - | - | - | T | - | V | K | - | L | - | - | - | K | - | - | - | - | - | - | - | - | - | - | - | R | - | - | L |   |

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L | V | T | A | D | H | S | H | V | F | S | F | G | G | Y | T | L | R | G | T | S | I | F | G | L | A | P | S | K | A | L |
| - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | S | - | - | - | - | - | - | - | - | Q |   |   |
| - | I | - | - | - | - | - | - | - | A | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | L | N | - | Q |
| I | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | L | N | - | - |

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| E | D | P | S | Y | Q | Q | Q | A | A | V | P | Q | A | S | E | T | H | G | G | E | D | V | A | V | F | A | R | G | P | Q |
| G | S | - | D | - | - | - | - | - | - | - | - | - | L | S | - | - | - | - | - | - | - | - | - | - | - | - | - | - |   |   |
| G | - | V | N | - | K | - | - | - | - | - | - | - | L | S | - | - | - | - | - | - | - | I | - | - | - | - | - | - |   |   |
| S | G | S | - | - | R | R | - | - | - | - | - | - | V | K | - | - | - | - | - | - | - | I | - | - | - | - | - | - |   |   |

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P | A | P | T | T | A | T | S | I | P | D | A | A | H | * | L | A | A | S | P | P | P | L | A | L | L | A | G | A | M | L |
| A | P | - | A | C | T | - | * | * | * | D | A | A | H | P | V | - | - | - | L | - | L | - | A | G | T | L | L | L | L | G |
| A | P | - | A | D | E | N | R | P | T | T | P | V | Q | N | S | - | I | T | M | N | N | V | - | - | S | L | Q | L | - | V |
| A | P | - | A | D | E | S | Q | T | T | T | T | T | R | Q | T | T | I | T | T | T | T | T | T | T | T | T | T | T | T | P |

FIG.3C

10/11

DKAYMALTEAGMFDNAIAKANE L T S E L D T L I 352 b.IAP  
GV--Q--V--D--ER--GQ--E--T 352 h.IAP  
GT--L--V--S--E--SQ--N--K--T 352 r.IAP  
GT--L--V--L--ER--SQ--R--T 352 m.IAP

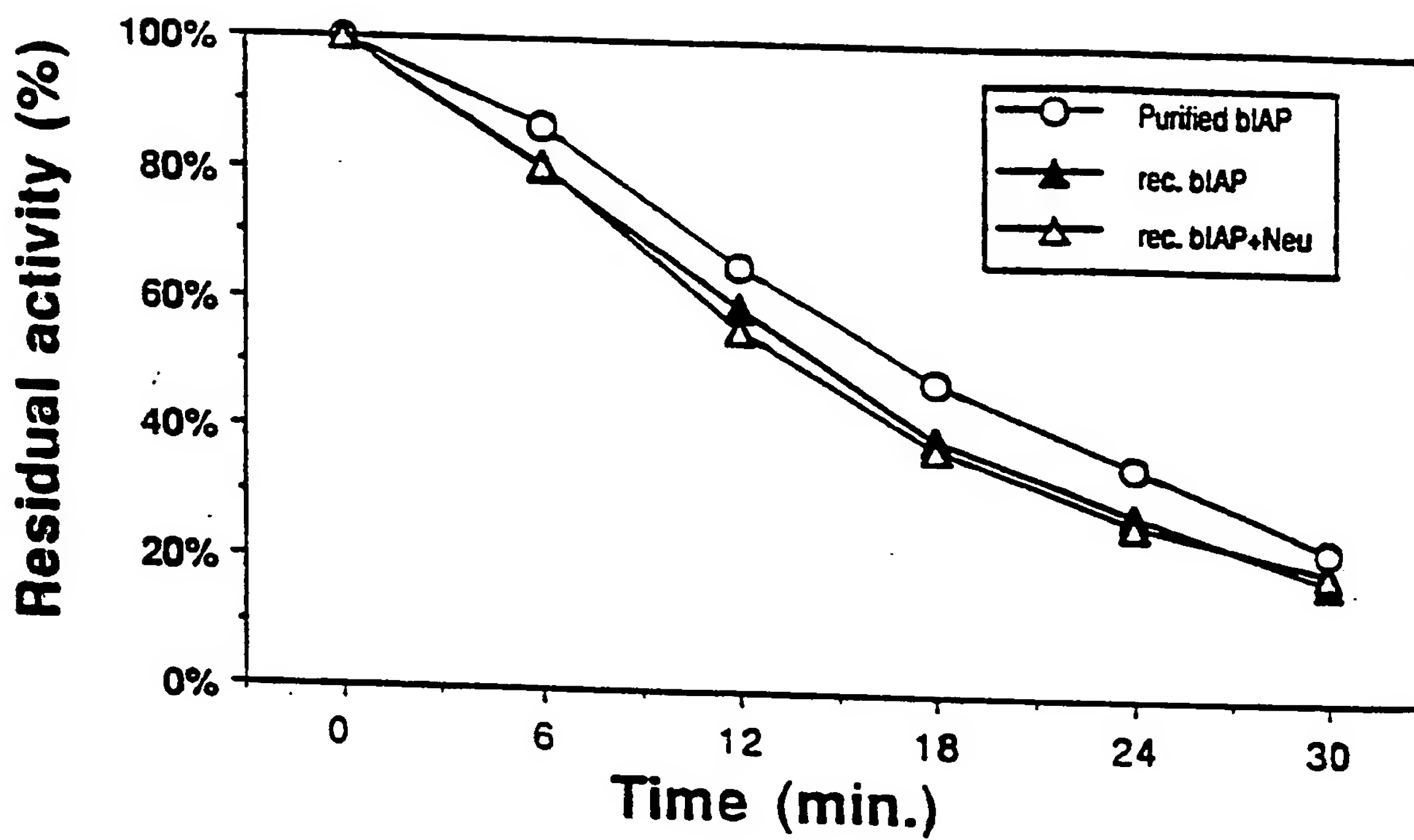
DSKSYTSLYGNPGYALGGSRPDVNDSTS 414 b.IAP  
--A--T--V--V--V--V--E--SE-- 414 h.IAP  
-G--V--V--V--V--N--N--T--AE-- 414 r.IAP  
-G-P--V--V--V--V--T--E--T--AAE-- 413 m.IAP

AHLVHGVEEETFVAHI MAFAGC VEPYTD CNL 476 b.IAP  
--Q--Q--Q--S--V--V--A--L--A--D-- 476 h.IAP  
--Q--Q--Q--N--Y--I--V--V--L--G-- 476 r.IAP  
--Q--Q--Q--N--Y--I--V--V--L--G-- 475 m.IAP

LLLA PTLY 514 b.IAP  
ASAAP 509 h.IAP  
SMLLVGTALVVS 520 r.IAP  
VHNSARSLGPATAPLALLAGMLLLGAPAE 540 m.IAP

FIG.3D

Figure 4



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US93/02172

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) : C12N 9/16, 15/00; C12Q 1/68, 1/42; A61K 39/00, 37/54

US CL : 435/196, 7.1, 6, 21, 172.3, 935/47, 424/85.8, 94.6

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/196, 7.1, 6, 21, 172.3, 935/47, 424/85.8, 94.6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|-----------|---|-----------------------|
| Y         | US, A, 4,707,438 (Keydar) 17 November 1987, see entire document.  | 1-20                  |
| Y         | PROQ. CLIN. BIOL. RES., Volume 344, issued 1990, J.L. Millan, "Oncodevelopmental Alkaline Phosphatases: In Search for a Function", pages 453-475, see especially page 460.                                      | 1-20                  |
| Y         | BIOCHEMICA ET BIOPHYSICA ACTA, Volume 831, issued 1985, J. Culp et al., "The active-site and amino-terminal amino acid sequence of bovine intestinal alkaline phosphatase", pages 330-334, see entire document. | 1-20                  |

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

|   |    |  |
|---|----|--|
| * Special categories of cited documents:  | *T | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  |
| "A" document defining the general state of the art which is not considered to be part of particular relevance   | *X | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone   |
| "E" earlier document published on or after the international filing date  | *Y | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | *& | document member of the same patent family  |
| "O" document referring to an oral disclosure, use, exhibition or other means  |    |  |
| "P" document published prior to the international filing date but later than the priority date claimed  |    |  |

Date of the actual completion of the international search

22 April 1993

Date of mailing of the international search report

04 MAY 1993

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Authorized officer

KEITH D. HENDRICKS

Facsimile No. NOT APPLICABLE

Telephone No. (703) 308-0196



# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US93/02172

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|-----------|---|-----------------------|
| Y         | GENOMICS, Volume 8, issued 1990, T. Manes et al., "Genomic structure and comparison of mouse tissue-specific alkaline phosphatase genes", pages 541-554.  | 1-20                  |
| Y         | THE JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 260, Number 20, issued 15 September 1985, M. Besman et al., "Isozymes of Bovine Intestinal Alkaline Phosphatase", pages 11190-11193, see entire document. | 1-20                  |



**THIS PAGE BLANK (USPTO)**